

Differential Privacy in Practice

Daniel Kifer

Boston University DP Summer School

Disclaimer

The views and opinions presented here are my own and do not represent the policy or official position of any government agency, company, or university.

- Privacy Semantics notebook:
<https://tinyurl.com/SemanticsNotebook>
- Utility experiments notebook: <https://tinyurl.com/UtilityNotebook>
- Make a copy, save to drive.

Take-home message #1



: “It’s a solved problem, there is a paper about it.”

Take-home message #1



: “It’s a solved problem, there is a paper about it.”

DETECT LANGUAGE ACADEMIC SPANISH FRENCH ▼ ↔ ENGLISH SPANISH ARABIC ▼

It's a solved problem, there is a paper about it. ×

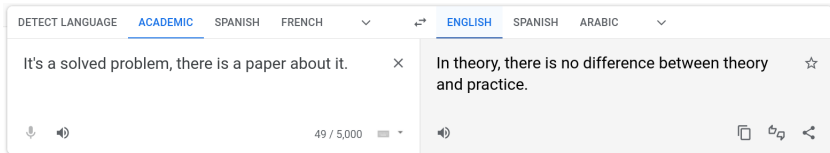
In theory, there is no difference between theory and practice. ☆

49 / 5,000

Take-home message #1



: “It’s a solved problem, there is a paper about it.”



: “But in practice there is.”

Take-home message # 2

- Differential privacy forces people to confront the unpleasant side of data analysis.
- Even without DP:
 - Can't perform arbitrary analyses on data 🧠🔒.
 - Data are always noisy.
 - **Other** people also have the right to privacy.
- DP forces re-examination of every aspect of data life cycle.

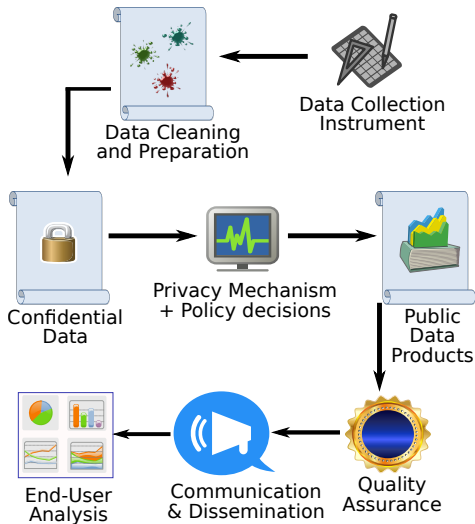
Main Focus

- Semantic privacy guarantees
- Utility issues in practice
- Also: other things to pay attention to

Outline

- 1 The Big Picture
- 2 Privacy Loss Random Variables
- 3 From PLRV to Privacy Definitions
- 4 Frequentist Guarantees
- 5 Bayesian Guarantees
- 6 Application to the TopDown Algorithm
- 7 Utility and Uncertainty

The Data Life Cycle



Data Collection

- Industry (often decentralized model)
 - Collection via deployed code
 - DP is not a substitute for informed consent!
- Statistical agencies (centralized model)
 - Surveys
 - Administrative records

Census Questionnaire

5. Please provide information for each person living here. If there is someone living here who pays the rent or owns this residence, start by listing him or her as Person 1. If the owner or the person who pays the rent does not live here, start by listing any adult living here as Person 1.

What is Person 1's name? *Print name below.*

First Name MI

Last Name(s)

6. What is Person 1's sex? Mark ☒ ONE box.

☐ Male ☐ Female

7. What is Person 1's age and what is Person 1's date of birth? For babies less than 1 year old, do not write the age in months. Write 0 as the age.

Age on April 1, 2020 Print numbers in boxes.

Month Day Year of birth

years

→ NOTE: Please answer BOTH Question 8 about Hispanic origin and Question 9 about race. For this census, Hispanic origins are not races.

8. Is Person 1 of Hispanic, Latino, or Spanish origin?

- ☐ No, not of Hispanic, Latino, or Spanish origin
- ☐ Yes, Mexican, Mexican Am., Chicano
- ☐ Yes, Puerto Rican
- ☐ Yes, Cuban
- ☐ Yes, another Hispanic, Latino, or Spanish origin – *Print, for example, Salvadoran, Dominican, Colombian, Guatemalan, Spaniard, Ecuadorian, etc.*

9. What is Person 1's race?

Mark ☒ one or more boxes **AND** print origins.

☐ White – *Print, for example, German, Irish, English, Italian, Lebanese, Egyptian, etc.*

☐ Black or African Am. – *Print, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc.*

☐ American Indian or Alaska Native – *Print name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow, Inupiat Traditional Government, Nome Eskimo Community, etc.*

☐ Chinese ☒ Vietnamese ☐ Native Hawaiian

☒ Filipino ☐ Korean ☐ Samoan

☐ Asian Indian ☐ Japanese ☐ Chamorro

☐ Other Asian – *Print, for example, Pakistani, Cambodian, Hmong, etc.* ☐ Other Pacific Islander – *Print, for example, Tongan, Fijian, Marshallese, etc.*

☐ Some other race – *Print race or origin.*

- Implications?
- Multiple Choice
- Fill-in the blank
- Redundancies

Data Cleaning and Preparation

- Automated collection:
 - Data poisoning
 - Hardware/software anomalies
- Self-reported data:
 - Often wrong
 - Sensitive to question phrasing
 - Missing data
- Administrative data:
 - Data entry errors
 - Out of date
- Important Questions:
 - How do you detect data anomalies?
 - How do you fix them?
 - How do you handle missing values?
 - ... in a manner consistent with differential privacy?

Examples

- 1 Birth date is provided, Age is missing

Examples

- ① Birth date is provided, Age is missing
 - Compute Age from birth date, no privacy impact

Examples

- 1 Birth date is provided, Age is missing
- 2 Entire record is missing

Examples

- ① Birth date is provided, Age is missing
- ② Entire record is missing
 - Hot-deck imputation
 - Copy over someone else's record (e.g., neighbor)
 - Privacy impact: neighbor affects 2 (or more) records

Examples

- 1 Birth date is provided, Age is missing
- 2 Entire record is missing
- 3 Race is missing

Examples

- ① Birth date is provided, Age is missing
- ② Entire record is missing
- ③ Race is missing
 - Build predictive model for race using the data
 - or find a similar (donor) record with race filled in (hot deck)
 - Impute race
 - Privacy impact: 1 person affects more than 1 record.
 - Differentially private imputation?

Examples

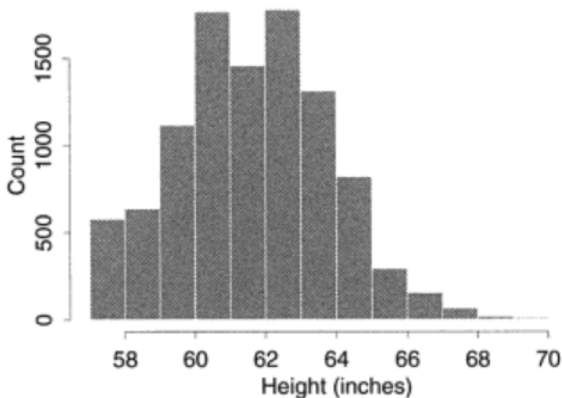
- ① Birth date is provided, Age is missing
- ② Entire record is missing
- ③ Race is missing
- ④ Yearly income placed in monthly income field.
 - Need to detect anomalies
 - Need to verify and fix them (statistical editing)

Challenges of Editing

- Many statistical editing rules come from examining the data.
 - E.g., many records with children older than parents.
 - Fix is record dependent:
 - Swap ages of children and parents?
 - Adjust ages of children to be \leq parents?
 - Mark them as non-children?
 - Fix may depend on aggregate properties of the data (privacy impact)?
- Big open question for DP, not much research.

Data quality

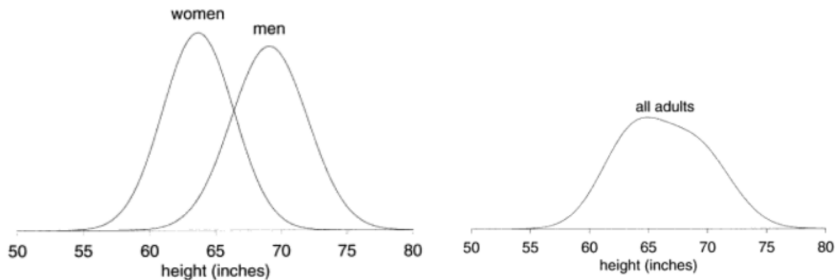
- Height of Conscripts in 1850's in Doubs, France
- Quality issue (beyond individual record repair) or actual property of population?



Gelman & Nolan, "Teaching Statistics: a bag of tricks" 2002

Hypothesis 1: Two separate populations

- Combining two populations
- Heights of men and women in U.S.



Gelman & Nolan, "Teaching Statistics: a bag of tricks" 2002

Doubs Data

- Data was measured to nearest centimeter
- Converted to inches
- Rounded to nearest inch

cm range	cm (est)	inches (est)	inch range
152 – 153	152.5	60.04	60–61
153 – 154	153.5	60.40	60–61
154 – 155	154.5	60.83	60–61
155 – 156	155.5	61.22	61–62
156 – 157	156.5	61.61	61–62
157 – 158	157.5	62.01	62–63
158 – 159	158.5	62.40	62–63
159 – 160	159.5	62.80	62–63
160 – 161	160.5	63.19	63–64
161 – 162	161.5	63.58	63–64
162 – 163	162.5	63.98	63–64

Gelman & Nolan, “Teaching Statistics: a bag of tricks” 2002

Privacy Mechanisms and Policy Decisions

- Privacy mechanisms are well studied and many exist.

Privacy Mechanisms and Policy Decisions

- Privacy mechanisms are well studied and many exist.
- Randomness
 - Do not use default random number generator
 - Census paper [Garfinkel and Leclerc, 2020]
 - Spark & others might duplicate random numbers

Privacy Mechanisms and Policy Decisions

- Privacy mechanisms are well studied and many exist.
- Randomness
 - Do not use default random number generator
 - Census paper [Garfinkel and Leclerc, 2020]
 - Spark & others might duplicate random numbers
- Floating point
 - “Continuous” noise distribution floating-point exploits [Mironov, 2012].
 - Some available options:
 - Secure exponential mechanism (base 2) [Ilvento, 2020]
 - Secure Discrete Laplace [Canonne et al., 2020]
 - Secure Discrete Gaussian [Canonne et al., 2020]

Privacy Mechanisms and Policy Decisions

- Privacy mechanisms are well studied and many exist.
- Randomness
 - Do not use default random number generator
 - Census paper [Garfinkel and Leclerc, 2020]
 - Spark & others might duplicate random numbers
- Floating point
 - “Continuous” noise distribution floating-point exploits [Mironov, 2012].
 - Some available options:
 - Secure exponential mechanism (base 2) [Ilvento, 2020]
 - Secure Discrete Laplace [Canonne et al., 2020]
 - Secure Discrete Gaussian [Canonne et al., 2020]
- Privacy Semantics
 - Analyze privacy impact of algorithmic choices (our focus)

Privacy Mechanisms and Policy Decisions

- Privacy mechanisms are well studied and many exist.
- Randomness
 - Do not use default random number generator
 - Census paper [Garfinkel and Leclerc, 2020]
 - Spark & others might duplicate random numbers
- Floating point
 - “Continuous” noise distribution floating-point exploits [Mironov, 2012].
 - Some available options:
 - Secure exponential mechanism (base 2) [Ilvento, 2020]
 - Secure Discrete Laplace [Canonne et al., 2020]
 - Secure Discrete Gaussian [Canonne et al., 2020]
- Privacy Semantics
 - Analyze privacy impact of algorithmic choices (our focus)
- Testing & Debugging (not research code).
 - Prevent data leakage
 - Test primitives, test privacy accounting
 - <https://www.nist.gov/blogs/cybersecurity-insights/testing-differential-privacy-bugs>
 - Minimize attack surface in code design

Public Data Products

- Data Format (our focus):
 - Privacy-Protected Microdata?
 - Tables?
 - Should entries be nonnegative?
 - Should entries be integers?
 - Should tables be consistent with each other?
 - Noisy Query Answers?
 - Each choice has important consequences.

Public Data Products

- Data Format (our focus):
 - Privacy-Protected Microdata?
 - Tables?
 - Should entries be nonnegative?
 - Should entries be integers?
 - Should tables be consistent with each other?
 - Noisy Query Answers?
 - Each choice has important consequences.
- No such thing as “general purpose data release”
 - Some analyses will be accurate, others not
 - Even without privacy protections, can't perform arbitrarily many statistical analyses.
- Stakeholders
 - Provide use-cases
 - Accuracy requirements
 - Often have competing interests for finite privacy loss budget

Quality Assurance

- Anomalies in the output data products (especially microdata):
 - Substantial errors.
 - Strange artifacts (e.g., too many regions where $\# \text{ male} = \# \text{ female}$)
 - Bugs that affect data quality but not privacy.
 - How to detect and fix (while maintaining privacy properties).
 - For 2020, U.S. Census Bureau used 2010 data for testing.

Communication and Dissemination

- Dissemination of data products.
- Dissemination of source code.
- Messaging strategy about benefits.
 - Many users don't perceive benefits as benefits.
 - Strong inertia for the status quo.

Example

- Decennial Census 2010 (a dramatic retelling):
 - We are going to modify the data (“data swapping”)
 - We are not going to draw attention to it.
 - You won’t get the details
 - You won’t know how much the data changed
 - You won’t have a chance to examine the effects.

Example

- Decennial Census 2010 (a dramatic retelling):
 - We are going to modify the data (“data swapping”)
 - We are not going to draw attention to it.
 - You won’t get the details
 - You won’t know how much the data changed
 - You won’t have a chance to examine the effects.
- Many users were fine with this

Dangers of Non-transparency [Alexander et al., 2010]

Inaccurate age and sex data in the Census PUMS files:

Evidence and Implications

J. Trent Alexander

Minnesota Population Center, University of Minnesota

Michael Davern

National Opinion Research Center, University of Chicago

Betsey Stevenson

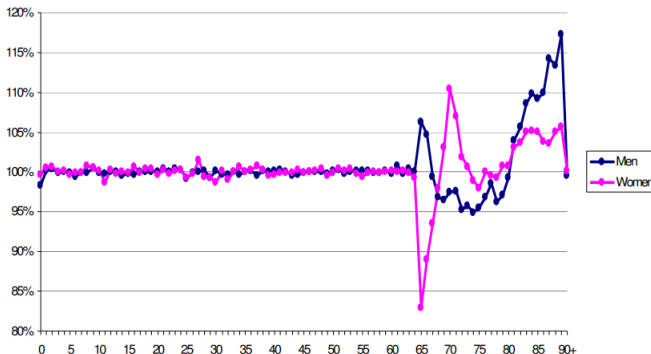
The Wharton School, University of Pennsylvania, CESifo, and NBER

Abstract

We discover and document errors in public use microdata samples ("PUMS files") of the 2000 Census, the 2003-2006 American Community Survey, and the 2004-2009 Current Population Survey. For women and men ages 65 and older, age- and sex-specific population estimates generated from the PUMS files differ by as much as 15% from counts in published data tables. Moreover, an analysis of labor force participation and marriage rates suggests the PUMS samples are not representative of the population at individual ages for those ages 65 and over. PUMS files substantially underestimate labor force participation of those near retirement ages and overestimate labor force participation rates of those at older ages. These problems were an unintentional by-product of the misapplication of a newer generation of disclosure avoidance procedures carried out on the data. The resulting errors in the public use data could significantly

Alexander, Davern, Stevenson, 2010

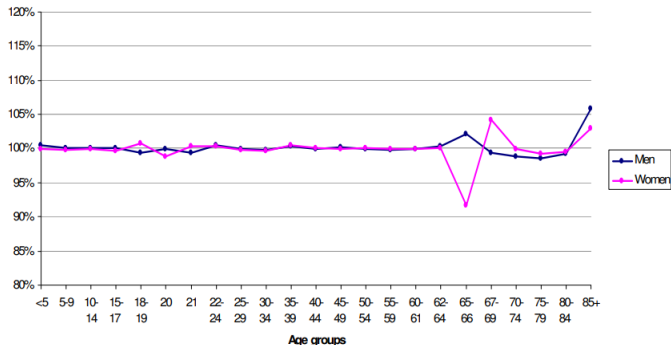
Figure 1. Population estimates from 2000 5% Census PUMS as a percentage of published data



Sources: Census 2000 Summary File 4, Table PCT3 (<http://factfinder.census.gov>); Census 2000 5% sample, IPUMS-USA (<http://usa.ipums.org/>).

Alexander, Davern, Stevenson, 2010

Figure 2. ACS 2006: Population estimates from PUMS as a percentage of published data



Sources: 2006 ACS Table B01001 (<http://factfinder.census.gov>); 2006 ACS PUMS, IPUMS-USA (<http://usa.ipums.org/>).

How was this detected?

- Study performed at Minnesota Population Center [Alexander et al., 2010]
- Pairs of publications: 2000 Census SF4 and PUMS, ACS PUMS and Tabulations, also CPS sex ratios.
 - They used inconsistent disclosure avoidance methods.
 - This is also a mistake.
 - Possibly leading to weaker privacy protections.
 - Discrepancy was larger than sampling error.
- Problem was corrected on the 2000 PUMS and ACS PUMS.
- It was never acknowledged or corrected on the CPS.
- Dictionary:
 - ACS: American Community Survey
 - PUMS: Public Use Microdata Samples
 - CPS: Current Population Survey
 - SF4: Summary File 4

Example

- Decennial Census 2020 (a dramatic retelling):
 - 2010 technology is broken.
 - We are going to modify the data.
 - You get **all** the details.
 - You **will** have a chance to examine the effects.
 - Your feedback will be used to improve the results.

Example

- Decennial Census 2020 (a dramatic retelling):
 - 2010 technology is broken.
 - We are going to modify the data.
 - You get all the details.
 - You will have a chance to examine the effects.
 - Your feedback will be used to improve the results.
- Many users were unenthusiastic

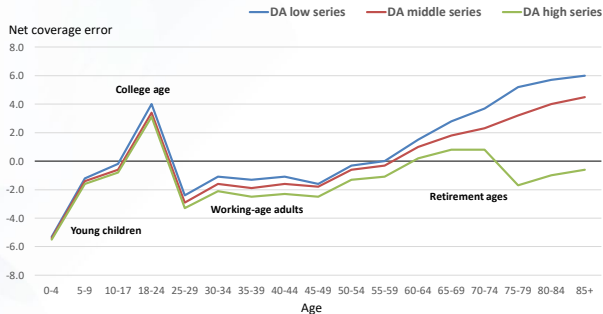
Example

- Decennial Census 2020 (a dramatic retelling):
 - 2010 technology is broken.
 - We are going to modify the data.
 - You get **all** the details.
 - You **will** have a chance to examine the effects.
 - Your feedback will be used to improve the results.
- Why did the messenger get shot?
 - Users are being forced to confront uncertainty.
 - But uncertainty has **always** been there (and was never negligible):
 - Differential net undercounts (enumeration differences by race, age, etc.)
 - Erroneous enumerations (location errors, duplicate responses)
 - Operational errors and gross omissions (nonresponse, frame errors)
 - Self-response errors (incomplete response)
 - Edit and imputation errors (including geolocation, characteristics)
 - Disclosure avoidance (swapping, synthetic responses)

Census 2020 PES DA News Conference

2020 Demographic Analysis (DA) Net Coverage Error Estimates for Selected Age Groups by Series: April 1, 2020

Post-Enumeration Survey and
Demographic Analysis

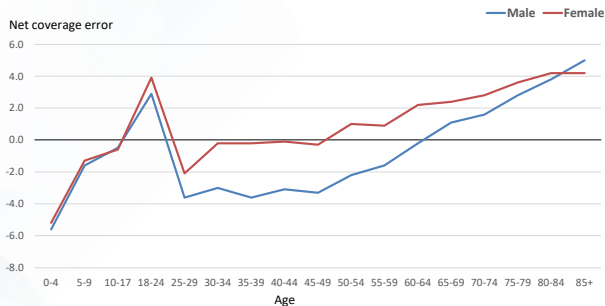


Source: U.S. Census Bureau, Population Division, 2020 Demographic Analysis (December 2020 release), and 2020 Census special tabulation (DRB Approval Number: CBDRB-FY22-DSEP-001).

Census 2020 PES DA News Conference

2020 Demographic Analysis Net Coverage Error Estimates for Selected Age Groups by Sex: April 1, 2020

Post-Enumeration Survey and
Demographic Analysis



Source: U.S. Census Bureau, Population Division, 2020 Demographic Analysis (December 2020 release), and 2020 Census special tabulation (DRB Approval Number: CBDRB-FY22-DSEP-001).

Census 2020 PES DA News Conference

Post-Enumeration Survey and
Demographic Analysis

Percent Net Coverage Error by Race and Hispanic Origin: 2010 and 2020

Race or Hispanic Origin	2010	2020
Total	0.01	-0.24
Race alone or in combination		
White	0.54*	0.66*
Non-Hispanic White alone	0.83*	1.64*
Black or African American	-2.06*	-3.30*
Asian	0.00	2.62*
American Indian or Alaska Native	-0.15	-0.91*
On Reservation	-4.88*	-5.64*
American Indian Areas Off Reservation	3.86	3.06
Balance of the United States	0.05	-0.86*
Native Hawaiian or Other Pacific Islander	-1.02	1.28
Some Other Race	-1.63*	-4.34*
Hispanic or Latino	-1.54*	-4.99*

*Net coverage error is statistically significantly different from 0.

Census 2020 PES DA News Conference

 Post-Enumeration Survey and
Demographic Analysis

Components of Census Coverage: 2010 and 2020 (In thousands)

Component of census coverage	2010				2020			
	Estimate	Standard error	Percent	Standard error	Estimate	Standard error	Percent	Standard error
Census count	300,700	X	100.0	X	323,200	X	100.0	X
Correct enumerations	284,700	199	94.7	0.07	305,100	145	94.4	0.04
Erroneous enumerations	10,040	199	3.3	0.07	7,167	145	2.2	0.04
Due to duplication	8,251	194	2.8	0.06	5,170	129	1.6	0.04
For other reasons	1,520	45	0.5	0.01	1,997	88	0.6	0.03
Whole-person imputations	5,993	X	2.0	X	10,850	X	3.4	X

X Not applicable.

2020 State Net Undercounts/Overcounts (Rounded)

State	Pop.	Error	%	90% Conf.
Arkansas	2,929,000	-184,000	-5.04	(-8.68%, -1.40%)
Delaware	967,000	53,123	5.45	(0.81%, 10.09%)
Florida	21,070,000	-733,000	-3.48	(-4.98%, -1.98%)
Hawaii	1,415,000	96,000	6.79	(4.03%, 9.55%)
Illinois	12,540,000	-247,000	-1.97	(-3.43%, -0.51%)
Massachusetts	6,784,000	152,000	2.24	(0.50%, 3.98%)
Minnesota	5,568,000	214,000	3.84	(2.24%, 5.44%)
Mississippi	2,868,000	-118,000	-4.11	(-6.79%, -1.43%)
New York	19,590,000	674,000	3.44	(1.89%, 4.99%)
Ohio	11,500,000	171,000	1.49	(0.39%, 2.59%)
Tennessee	6,754,000	-323,000	-4.78	(-7.26%, -2.30%)
Texas	28,540,000	-548,000	-1.92	(-3.27%, -0.57%)
Utah	3,216,000	83,000	2.59	(0.57%, 4.61%)

Differential Privacy vs. Undercount

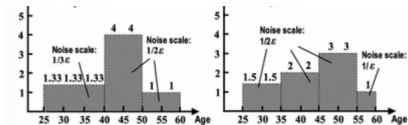
- Query:
 - Number of people in each state in AIAN areas.
 - Number of people in each state not in AIAN areas.
- Noise:
 - discrete gaussian [Canonne et al., 2020]
 - scale parameter: $\sigma^2 = 1.21$ (error is a few people)
- Texas:
 - Census population: 28,540,000
 - Estimated undercount: -548,000
 - Upper 90% undercount error estimate: -933,000
 - Lower 90% undercount error estimate: -163,000
- Caveats
 - Does not account for postprocessing in disclosure avoidance.
 - Estimated overcount/undercount by smaller geographies is not available.

Links

- Post-Enumeration Survey Demographic Analysis: <https://www.census.gov/content/dam/Census/newsroom/press-kits/2022/20220310-presentation-quality-news-conference.pdf>
- Coverage by State:
<https://www.census.gov/programs-surveys/decennial-census/about/coverage-measurement/pes.html>

End-user analysis

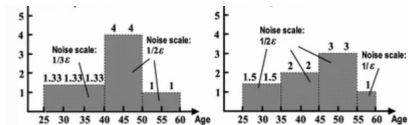
- First rule of statistics: don't underestimate variability in your data.
 - Leads to false discoveries
 - Must account for sources of uncertainty
- All disclosure avoidance algorithms might add properties to the data:
 - Compression style approaches can improve squared error
 - but add bias
 - data can appear to have less variability



[Xu et al., 2013]

End-user analysis

- First rule of statistics: don't underestimate variability in your data.
 - Leads to false discoveries
 - Must account for sources of uncertainty
- All disclosure avoidance algorithms might add properties to the data:
 - Compression style approaches can improve squared error
 - but add bias
 - data can appear to have less variability



[Xu et al., 2013]

- User must modify analysis to account for disclosure avoidance. Only ignorable in special situations [Abowd and Schmutte, 2015].
- No such thing as privacy-protected microdata that you can just use with your old software.
- Many end-users don't know what to do

Our Focus



- Privacy semantics.
 - May need to or want to customize privacy definitions.
 - Need to help decision-makers set privacy parameters.
 - Need to know how to evaluate privacy definitions.
 - Just because it “looks” like DP does not make it good.
- Unintended consequences of design choices.
 - Format of output data products
 - Choice of optimization criteria

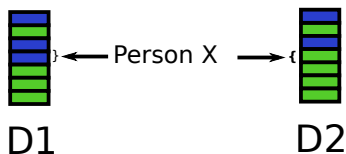
Outline

- 1 The Big Picture
- 2 Privacy Loss Random Variables
- 3 From PLRV to Privacy Definitions
- 4 Frequentist Guarantees
- 5 Bayesian Guarantees
- 6 Application to the TopDown Algorithm
- 7 Utility and Uncertainty

PLRV

- Fundamental concept behind differential privacy.
- Goals:
 - What it is
 - How to interpret it
 - How to compute/estimate it
 - Important properties
- Uses:
 - Create formal privacy definitions.
 - Interpret privacy guarantees.
 - Practical applications need tightest guarantees possible.

Neighbors



- Neighbors: specify what is being protected.
 - Bounded/change-a-record neighbors $\mathcal{D}_1, \mathcal{D}_2$
 - Can be used when participation in data is not sensitive
 - Contents of the records are sensitive
 - For simplicity, we use this here.
- Unbounded/add-remove neighbors:
 - Used to protect inference about participation
 - Used to protect record contents
 - Ideal version, use whenever possible

Randomized Response

- Mechanism M : a (randomized) algorithm designed to protect privacy.
- Differential privacy officially invented in 2006 [Dwork et al., 2006b]
- Mechanisms for differential privacy existed in 1965 [Warner, 1965]
- Face-to-face survey: “have you ever engaged in insider trading?”
 - Respondents are likely to lie
 - or withhold information

Randomized Response

- Mechanism M : a (randomized) algorithm designed to protect privacy.
- Differential privacy officially invented in 2006 [Dwork et al., 2006b]
- Mechanisms for differential privacy existed in 1965 [Warner, 1965]
- Face-to-face survey: “have you ever engaged in insider trading?”
 - Respondents are likely to lie
 - or withhold information
- Warner's Spinner:
 - Only the respondent sees spinner.
 - $P(\text{True}) = p > \frac{1}{2}$
 - If arrow lands on "True", answer truthfully
 - If arrow lands on "False", lie
 - Note: mechanism is public, randomness is not.



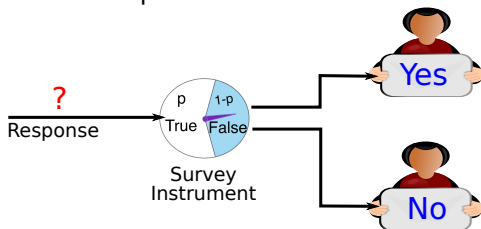
Randomized Response

- Mechanism M : a (randomized) algorithm designed to protect privacy.
- Differential privacy officially invented in 2006 [Dwork et al., 2006b]
- Mechanisms for differential privacy existed in 1965 [Warner, 1965]
- Face-to-face survey: “have you ever engaged in insider trading?”
 - Respondents are likely to lie
 - or withhold information
- Warner’s Spinner:
 - Only the respondent sees spinner.
 - $P(\text{True}) = p > \frac{1}{2}$
 - If arrow lands on “True”, answer truthfully
 - If arrow lands on “False”, lie
 - Note: mechanism is public, randomness is not.
- Two possible **ordered** pairs of neighboring datasets:
 - \mathcal{D}_1 : true answer is yes
 - \mathcal{D}_2 : true answer is no
 - Ordered pairs $(\mathcal{D}_1, \mathcal{D}_2)$ and $(\mathcal{D}_2, \mathcal{D}_1)$



What are the Odds?

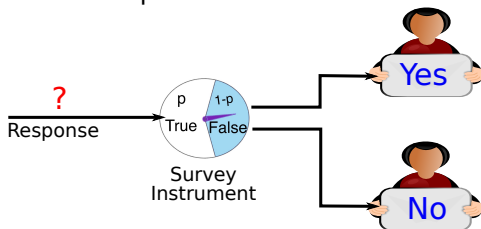
- Attacker's point of view:



- \mathcal{D}_1 : input response is yes
- \mathcal{D}_2 : input response is no
- $p = \frac{e^2}{1+e^2} \approx 0.88$
- log odds for output ω : $\log \frac{P(M(\mathcal{D}_1)=\omega)}{P(M(\mathcal{D}_2)=\omega)}$
 - log odds for $\omega = \text{"Yes"}$? _____
 - log odds for $\omega = \text{"No"}$? _____

What are the Odds?

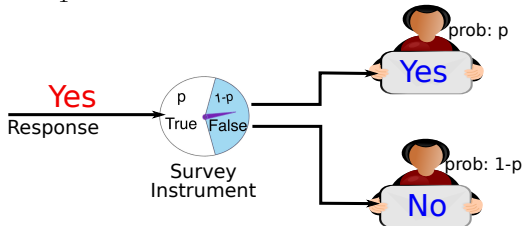
- Attacker's point of view:



- \mathcal{D}_1 : input response is yes
- \mathcal{D}_2 : input response is no
- $p = \frac{e^2}{1+e^2} \approx 0.88$
- log odds for output ω : $\log \frac{P(M(\mathcal{D}_1)=\omega)}{P(M(\mathcal{D}_2)=\omega)}$
 - log odds for $\omega = \text{"Yes"}$: 2
 - log odds for $\omega = \text{"No"}$: -2
- ω is random so odds are also random.

Privacy loss random variable

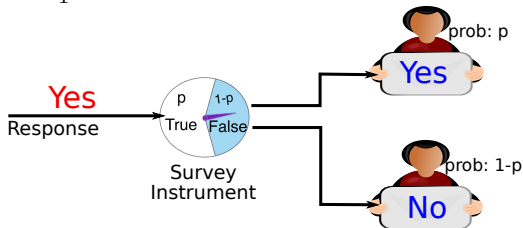
- If \mathcal{D}_1 is the true dataset:



- $p = \frac{e^2}{1+e^2} \approx 0.88$
- Probability of $\omega = \text{"Yes"}$
 - Probability: $\frac{e^2}{1+e^2}$ (a.k.a $P(M(\mathcal{D}_1) = \text{Yes})$)
 - Log odds (privacy loss): 2
 - Evidence for attacker to think \mathcal{D}_1 is true.
- Probability of $\omega = \text{"No"}$
 - Probability: $\frac{1}{1+e^2}$ (a.k.a $P(M(\mathcal{D}_1) = \text{No})$)
 - Log odds (privacy loss): -2
 - Evidence for attacker to think \mathcal{D}_1 is **not** true.

Privacy loss random variable

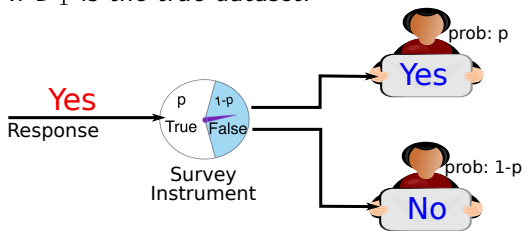
- If \mathcal{D}_1 is the true dataset:



- $p = \frac{e^2}{1+e^2} \approx 0.88$
- Privacy loss is 2 with probability $\frac{e^2}{1+e^2}$
- Privacy loss is -2 with probability $\frac{1}{1+e^2}$
- This is the privacy loss random variable $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$
 - Distribution of privacy loss when \mathcal{D}_1 is true.
 - Distribution of $\log \frac{P(M(\mathcal{D}_1)=\omega)}{P(M(\mathcal{D}_2)=\omega)}$ when $\omega \sim M(\mathcal{D}_1)$.

Privacy loss random variable (summary)

- If \mathcal{D}_1 is the true dataset:

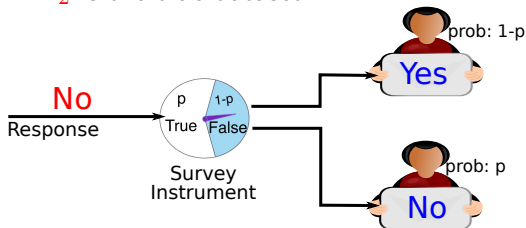


- $p = \frac{e^\epsilon}{1+e^\epsilon}$

$$\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M = \begin{cases} \epsilon & \text{with probability } \frac{e^\epsilon}{1+e^\epsilon} \\ -\epsilon & \text{with probability } \frac{1}{1+e^\epsilon} \end{cases}$$

Reversed PLRV

- PLRVs come in pairs.
- What about $\mathcal{L}_{\mathcal{D}_2, \mathcal{D}_1}^M$?
- If \mathcal{D}_2 is the true dataset:



- $p = \frac{e^\epsilon}{1+e^\epsilon}$

$$\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M = \begin{cases} ? \\ ? \end{cases}$$

Generalization


- \mathcal{D} is responses from n different people: $\{r_1, \dots, r_n\}$
- apply randomized response to each of them
- $\omega = \{\omega_1, \dots, \omega_n\}$
- Suppose \mathcal{D}_1 and \mathcal{D}_2 differ on Jessie.
- What is $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$?

$$\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M = \begin{cases} ? \\ ? \end{cases}$$

Jupyter Notebook



PLRV for the Geometric Mechanism

- Query $q(\mathcal{D}) = \#$ of “Yes” values.
- Sensitivity: changing one person’s record changes query answer by at most ± 1 .
- Geometric Mechanism:
 - $M(\mathcal{D}) = q(\mathcal{D}) + k$
 - where $P(k) = \frac{1-e^{-\epsilon}}{1+e^{-\epsilon}} e^{-|k|\epsilon}$
 - k is integer
- If $q(\mathcal{D}_1) = 6$ and $q(\mathcal{D}_2) = 5$, what is $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$?
(use Jupyter  or math)

Lessons Learned

- Computing or estimating $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ accurately is key to getting tightest privacy semantics in practice.
- Two non-equivalent mechanisms can have the same $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ for all pairs of neighbors.

Group privacy loss

- Neighbors of neighbors
 - $\mathcal{D}_1 = (0, 0)$, responses from Jessie and Alice
 - $\mathcal{D}_2 = (1, 1)$
 - How well are Jessie's and Alice's responses protected **together**?
 - Group privacy loss random variable $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$

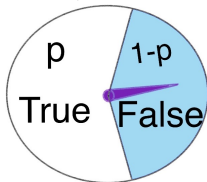
Group privacy loss

- Neighbors of neighbors

- $\mathcal{D}_1 = (0, 0)$, responses from Jessie and Alice
- $\mathcal{D}_2 = (1, 1)$
- How well are Jessie's and Alice's responses protected **together**?
- Group privacy loss random variable $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$

- Mechanism M_1 :

- parallel randomized response
- $p = \frac{e^\epsilon}{1+e^\epsilon}$



- Mechanism M_2 :

- geometric mechanism
- $q(\mathcal{D}) = \text{sum}(\mathcal{D})$
- $M(\mathcal{D}) = q(\mathcal{D}) + k$
- where $P(k) = \frac{1-e^{-\epsilon}}{1+e^{-\epsilon}} e^{-|k|\epsilon}$
- k is integer

What can we say about the group privacy loss random variable? 🌐


Lessons Learned

- Computing or estimating $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ accurately is key to getting tightest privacy semantics in practice.
- Two non-equivalent mechanisms can have the same $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ for all pairs of neighbors.
 - Group privacy loss random variables can tell them apart.


Gaussian Mechanism

- Used for queries q whose answers may not be integers.
- $M(\mathcal{D}) = q(\mathcal{D}) + N(0, \sigma^2)$
 - What is $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$? 🌐

Gaussian Mechanism

- Used for queries q whose answers may not be integers.
- $M(\mathcal{D}) = q(\mathcal{D}) + N(0, \sigma^2)$
 - What is $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$? 
- In practice, use discrete Gaussian [Canonne et al., 2020]
 - $M(\mathcal{D}) = \text{round}(q(\mathcal{D})) + N_Z(0, \sigma^2)$.
 - $N_Z(0, \sigma^2)$ pmf is $P(k) = \frac{e^{-k^2/(2\sigma^2)}}{\sum_j e^{-j^2/(2\sigma^2)}}$
 - Or round to some other discrete set.
 - PLRV is not so easy to work with (yet).

Laplace Mechanism

- Used for queries q whose answers may not be integers.
- $M(\mathcal{D}) = q(\mathcal{D}) + \text{Lap}(1/\epsilon)$
- $\text{Lap}(1/\epsilon)$ density $f(x) = \frac{\epsilon}{2} e^{-\epsilon|x|}$
 - What is $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$? 
- In practice, use Geometric Mechanism [Canonne et al., 2020]
 - $M(\mathcal{D}) = \text{round}(q(\mathcal{D})) + k.$
 - where $P(k) = \frac{1-e^{-\epsilon}}{1+e^{-\epsilon}} e^{-|k|\epsilon}$
 - Or round to some other discrete set.

Independent Composition

- Mechanism M_1 with $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M_1}$
- Mechanism M_2 with $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M_2}$
- Mechanism M^* runs both of them and returns their outputs.
 - $M^*(\mathcal{D}) = (M_1(\mathcal{D}), M_2(\mathcal{D}))$
 - What is the PLRV of M^* ? $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M^*} = ?$ 🌐

Independent Composition

- Mechanism M_1 with $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M_1}$
- Mechanism M_2 with $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M_2}$
- Mechanism M^* runs both of them and returns their outputs.
 - $M^*(\mathcal{D}) = (M_1(\mathcal{D}), M_2(\mathcal{D}))$
 - What is the PLRV of M^* ? $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M^*} = ?$ 🌐
- $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M^*} = \mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M_1} + \mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M_2}$
- Independent composition: M_1 and M_2 don't know about each other.

Lessons Learned

- Computing or estimating $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ accurately is key to getting tightest privacy semantics in practice.
- Two non-equivalent mechanisms can have the same $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ for all pairs of neighbors.
 - Group privacy loss random variables can tell them apart.
- If M_1 and M_2 are independently applied to the same data, the resulting PLRV is $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M_1} + \mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M_2}$
- Tail inequalities of PLRV are unstable: they can increase or decrease with postprocessing.

Outline

- 1 The Big Picture
- 2 Privacy Loss Random Variables
- 3 From PLRV to Privacy Definitions
- 4 Frequentist Guarantees
- 5 Bayesian Guarantees
- 6 Application to the TopDown Algorithm
- 7 Utility and Uncertainty

PLRVs and Privacy Definitions

- Each mechanism M has infinitely many PLRVs
 - $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ for every pair of neighbors $\mathcal{D}_1, \mathcal{D}_2$.
 - Also possible to consider $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ for non-neighbors.
- Represents distribution of log odds (privacy loss)
- Together, PLRVs capture privacy properties of M in the differential privacy framework.

PLRVs and Privacy Definitions

- Each mechanism M has infinitely many PLRVs
 - $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ for every pair of neighbors $\mathcal{D}_1, \mathcal{D}_2$.
 - Also possible to consider $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ for non-neighbors.
- Represents distribution of log odds (privacy loss)
- Together, PLRVs capture privacy properties of M in the differential privacy framework.
 - Unwieldy to work with infinitely many distributions.
 - Difficult to explain to policy makers and public, or any non-expert.
 - Not all properties of PLRVs are relevant to privacy.
- Privacy definitions: required properties of PLRVs.
- Privacy accounting frameworks: techniques for computing/keeping track of these properties.

Example 1: Pure Differential Privacy

Definition (Pure ϵ -Differential Privacy [Dwork et al., 2006b])

Given $\epsilon \geq 0$, M satisfies pure DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$,

- For all ω , $P(M(\mathcal{D}_1) = \omega) \leq e^\epsilon P(M(\mathcal{D}_2) = \omega)$
- Or, equivalently, $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \leq \epsilon$ with probability 1.

*Using discrete notation for convenience.

Example 1: Pure Differential Privacy

Definition (Pure ϵ -Differential Privacy [Dwork et al., 2006b])

Given $\epsilon \geq 0$, M satisfies pure DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$,

- For all ω , $P(M(\mathcal{D}_1) = \omega) \leq e^\epsilon P(M(\mathcal{D}_2) = \omega)$
- Or, equivalently, $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \leq \epsilon$ with probability 1.

*Using discrete notation for convenience.

- Pros:
 - Intuitive privacy parameter (upper bound on odds).
 - Easy to work with.
 - Many results in literature.

Example 1: Pure Differential Privacy

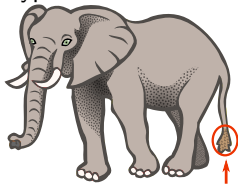
Definition (Pure ϵ -Differential Privacy [Dwork et al., 2006b])

Given $\epsilon \geq 0$, M satisfies pure DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$,



- For all ω , $P(M(\mathcal{D}_1) = \omega) \leq e^\epsilon P(M(\mathcal{D}_2) = \omega)$
- Or, equivalently, $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \leq \epsilon$ with probability 1.

*Using discrete notation for convenience.

- Cons:
 - Incomplete picture of privacy properties
 - Typical PLRV behavior may not be accurately described by max.



Important Properties

- Transparency:
 - Algorithm is public
 - Does not affect privacy guarantees
 - Public can check for data quality bugs that might not be caught otherwise.
-  Be prepared for criticism.
 - Many papers criticize effects of DP algorithms (e.g., Census data) because it is transparent.
 - Missing comparisons to data quality issues:
 - Estimated 5% undercount of Hispanic or Latino
 - Estimated net undercount in Texas of half million people
 - Estimates 10 million whole person imputations
 - Broad patterns are known, exact details are not.
 - Very few similar comparisons to data swapping [Christ et al., 2022]
- Privacy definitions based on $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ are transparent .

Postprocessing Invariance

- If M satisfies a privacy definition at specific parameter values (e.g., ϵ) so should $A \circ M$.
 - (range of M is in domain of A)
 - A could create tables.
 - A could create a population forecasting model.
 - A could combine the output with other data to recommend policy actions.
 - A could try to determine if Kobbi is TCS+.
- If $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ satisfies privacy conditions, so should $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{A \circ M}$

Postprocessing Invariance

- If M satisfies a privacy definition at specific parameter values (e.g., ϵ) so should $A \circ M$.
 - (range of M is in domain of A)
 - A could create tables.
 - A could create a population forecasting model.
 - A could combine the output with other data to recommend policy actions.
 - A could try to determine if Kobbi is TCS+.
- If $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ satisfies privacy conditions, so should $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{A \circ M}$
- Why is it important?
 - Privacy-protected data are postprocessed, published, repackaged.
 - Post-processing invariance means you accounted for subsequent uses when evaluating privacy protections.
 - Some DP variations are not post-processing invariant. ⚠
 - No postprocessing invariance \Rightarrow no adaptive composition.

Composition 1

- Independent Composition
- Mechanisms M_1, \dots, M_k
 - Mechanisms chosen in advance
 - Mechanisms are unaware of each other.
- Privacy properties of each M_i are known in isolation.
 - E.g., M_i satisfies pure ϵ_i differential privacy.
- Combined privacy property of releasing $M_1(\mathcal{D}), \dots, M_k(\mathcal{D})$?
 - E.g, $(\sum_i \epsilon_i)$ -differential privacy.

Composition 1

- Independent Composition
- Mechanisms M_1, \dots, M_k
 - Mechanisms chosen in advance
 - Mechanisms are unaware of each other.
- Privacy properties of each M_i are known in isolation.
 - E.g., M_i satisfies pure ϵ_i differential privacy.
- Combined privacy property of releasing $M_1(\mathcal{D}), \dots, M_k(\mathcal{D})$?
 - E.g, $(\sum_i \epsilon_i)$ -differential privacy.
- $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M_1, \dots, M_k} = \mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M_1} + \dots + \mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{M_k}$
- Requirements:
 - Minimum: releasing $M_1(\mathcal{D}), \dots, M_k(\mathcal{D})$ should not have better properties than releasing only $M_1(\mathcal{D})$
 - Ideally: composition should be (sub)linear

Why Composition is Important

- **Independent Composition:**
 - Simplifies algorithm construction.
 - Identify privacy properties of each component, obtain overall privacy properties.
 - Predictable degradation in privacy over multiple releases.
 - Redistricting data
 - Demographics data

Composition 2

- Adaptive Composition
- “If there are enough people in a Census block, I want to know the ages and races, if not, then just the races”

M_1 : number of people in the block + noise

if *noisy people* $\geq N$ **then**

| M_{2a} : noisy Race histogram and noisy Age histogram

else

| M_{2b} : noisy Race histogram

end

- Result of M_1 influences choice of M_2 , etc.

Composition 2

- Adaptive Composition
- Privacy budgets chosen in advance (e.g., $\epsilon_1, \dots, \epsilon_k$).
- Mechanisms are history aware:
 - $M_1(\mathcal{D})$ produces ω_1
 - $M_2(\omega_1, \mathcal{D})$ produces ω_2
 - $M_3(\omega_2, \mathcal{D})$ produces ω_3 .
 - etc.
- Privacy:
 - M_1 satisfies privacy (e.g., ϵ_1 -DP).
 - For each ω_1 , $M_2(\omega_1, \cdot)$ satisfies privacy (e.g., ϵ_2 -DP)
 - For each ω_2 , $M_3(\omega_2, \cdot)$ satisfies privacy (e.g., ϵ_3 -DP)
- Combined privacy property of releasing $\omega_1, \dots, \omega_k$?
 - E.g, $(\sum_i \epsilon_i)$ -differential privacy.
- Requirements:
 - Minimum: releasing $\omega_1, \dots, \omega_k$ should not have better properties than releasing only ω_1 .
 - Ideally: composition should be (sub)linear

Why Composition is Important

- Independent Composition:
 - Simplifies algorithm construction.
 - Identify privacy properties of each component, obtain overall privacy properties.
 - Predictable degradation in privacy over multiple releases.
 - Redistricting data
 - Demographics data
- Adaptive Composition:
 - Wider variety of mechanisms is possible.
 - Adaptive mechanisms are not always preferable.
 - Future data releases can depend on current data release.

Composition 3

- Fully Adaptive Composition
- Privacy budget for **first** mechanism is chosen.
- Mechanisms are history aware:
 - $M_1(\mathcal{D})$ outputs ω_1 and privacy parameters (e.g., ϵ_2).
 - $M_2(\omega_1, \epsilon_2, \mathcal{D})$ produces ω_2 and privacy params (e.g., ϵ_3).
 - $M_3(\omega_2, \epsilon_3, \mathcal{D})$ produces ω_3 and privacy params.
 - etc.
- Privacy:
 - M_1 satisfies privacy (e.g., ϵ_1 -DP).
 - For each ω_1 and ϵ_2 , $M_2(\omega_1, \epsilon_2, \cdot)$ satisfies privacy (e.g., ϵ_2 -DP)
 - For each ω_2 and ϵ_3 , $M_3(\omega_2, \epsilon_3, \cdot)$ satisfies privacy (e.g., ϵ_3 -DP)
- Combined privacy property of releasing $\omega_1, \dots, \omega_k$?
 - E.g, ϵ -differential privacy when $\sum_i \epsilon_i \leq \epsilon$ with probability 1.
- Requirements:
 - Minimum: releasing $\omega_1, \dots, \omega_k$ should not have better properties than releasing only ω_1 .
 - Ideally: composition should be (sub)linear

Why Composition is Important

- Independent Composition:

- Simplifies algorithm construction.
- Identify privacy properties of each component, obtain overall privacy properties.
- Predictable degradation in privacy over multiple releases.
 - Redistricting data
 - Demographics data

- Adaptive Composition:

- Wider variety of mechanisms is possible.
- Adaptive mechanisms are not always preferable.
- Future data releases can depend on current data release.

- Full Adaptive Composition:

- Even wider variety of mechanisms is possible.
- Adaptive mechanisms are not always preferable.
- Privacy budget allocation decisions for data releases can be made on the fly.

Convexity

- M_1 satisfies a privacy definition.
- M_2 satisfies same privacy definition (with same parameters).
- $M_p(\mathcal{D}) = \begin{cases} M_1(\mathcal{D}) & \text{with probability } p \\ M_2(\mathcal{D}) & \text{otherwise} \end{cases}$
 - M_p should also satisfy same privacy definition.
 - Useful sanity check for privacy definitions.

Convexity

- M_1 satisfies a privacy definition.
- M_2 satisfies same privacy definition (with same parameters).
- $M_p(\mathcal{D}) = \begin{cases} M_1(\mathcal{D}) & \text{with probability } p \\ M_2(\mathcal{D}) & \text{otherwise} \end{cases}$
 - M_p should also satisfy same privacy definition.
 - Useful sanity check for privacy definitions.
- Weak form of adaptive composition:
 - $M'(\mathcal{D})$ ignores \mathcal{D} and outputs a biased bit.
 - $M''(\mathcal{D})$ uses biased bit to determine if it runs M_1 or M_2 .
 - Total privacy cost equals cost of M_1 (= cost of M_2)

Convexity

- M_1 satisfies a privacy definition.
- M_2 satisfies same privacy definition (with same parameters).
- $M_p(\mathcal{D}) = \begin{cases} M_1(\mathcal{D}) & \text{with probability } p \\ M_2(\mathcal{D}) & \text{otherwise} \end{cases}$
 - M_p should also satisfy same privacy definition.
 - Useful sanity check for privacy definitions.
- Weak form of adaptive composition:
 - $M'(\mathcal{D})$ ignores \mathcal{D} and outputs a biased bit.
 - $M''(\mathcal{D})$ uses biased bit to determine if it runs M_1 or M_2 .
 - Total privacy cost equals cost of M_1 (= cost of M_2)
- **What convexity is not:**
 - If M_1 and M_2 have different privacy parameters.
 - Convexity says nothing.
 - Privacy parameters of M_p are not an average of M_1 and M_2 .
 - If M_1 is 1-DP and M_2 is 2-DP, then M_p is 2-DP (not an average of 1 and 2).

Back to Privacy Definitions

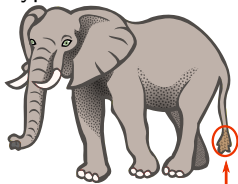
Definition (Pure ϵ -Differential Privacy [Dwork et al., 2006b])

Given $\epsilon \geq 0$, M satisfies pure DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$,

- For all ω , $P(M(\mathcal{D}_1) = \omega) \leq e^\epsilon P(M(\mathcal{D}_2) = \omega)$
- Or, equivalently, $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \leq \epsilon$ with probability 1.

*Using discrete notation for convenience.

- Cons:
 - Incomplete picture of privacy properties
 - Typical PLRV behavior may not be accurately described by max.



Example 2: Approximate Differential Privacy

Definition (Approximate Differential Privacy [Dwork et al., 2006a])

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies (ϵ, δ) -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$

- For all measurable sets S , $P(M(\mathcal{D}_1) \in S) \leq e^\epsilon P(M(\mathcal{D}_2) \in S) + \delta$
- Or, equivalently, $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \geq \epsilon) - e^\epsilon P(\mathcal{L}_{\mathcal{D}_2, \mathcal{D}_1}^M \leq -\epsilon) \leq \delta$.

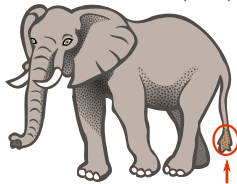
Example 2: Approximate Differential Privacy

Definition (Approximate Differential Privacy [Dwork et al., 2006a])

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies (ϵ, δ) -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$

- For all measurable sets S , $P(M(\mathcal{D}_1) \in S) \leq e^\epsilon P(M(\mathcal{D}_2) \in S) + \delta$
- Or, equivalently, $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \geq \epsilon) - e^\epsilon P(\mathcal{L}_{\mathcal{D}_2, \mathcal{D}_1}^M \leq -\epsilon) \leq \delta$.

- Single pair of (ϵ, δ) -parameters:



- But M satisfies a continuum of (ϵ, δ) -parameters.

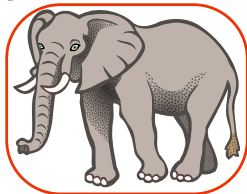
Example 2: Approximate Differential Privacy

Definition (Approximate Differential Privacy [Dwork et al., 2006a])

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies (ϵ, δ) -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$

- For all measurable sets S , $P(M(\mathcal{D}_1) \in S) \leq e^\epsilon P(M(\mathcal{D}_2) \in S) + \delta$
- Or, equivalently, $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \geq \epsilon) - e^\epsilon P(\mathcal{L}_{\mathcal{D}_2, \mathcal{D}_1}^M \leq -\epsilon) \leq \delta$.

- (ϵ, δ) -curve captures the distribution of the PLRV
[Dong et al., 2022, Sommer et al., 2019]

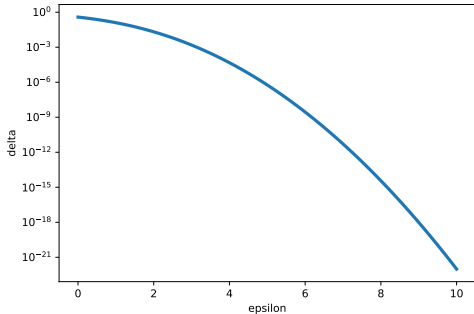


Example 2: Approximate Differential Privacy

Definition (Approximate Differential Privacy [Dwork et al., 2006a])

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies (ϵ, δ) -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$

- For all measurable sets S , $P(M(\mathcal{D}_1) \in S) \leq e^\epsilon P(M(\mathcal{D}_2) \in S) + \delta$
- Or, equivalently, $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \geq \epsilon) - e^\epsilon P(\mathcal{L}_{\mathcal{D}_2, \mathcal{D}_1}^M \leq -\epsilon) \leq \delta$.
- $M(\mathcal{D}) = \# \text{ AIAN in } \mathcal{D} + N(0, 1)$.



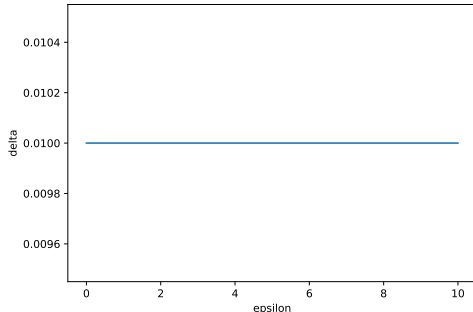
Example 2: Approximate Differential Privacy

Definition (Approximate Differential Privacy [Dwork et al., 2006a])

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies (ϵ, δ) -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$

- For all measurable sets S , $P(M(\mathcal{D}_1) \in S) \leq e^\epsilon P(M(\mathcal{D}_2) \in S) + \delta$
- Or, equivalently, $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \geq \epsilon) - e^\epsilon P(\mathcal{L}_{\mathcal{D}_2, \mathcal{D}_1}^M \leq -\epsilon) \leq \delta$.

- $M(\mathcal{D})$ = randomly chosen record in dataset of size 100.




Example 2: Approximate Differential Privacy

Definition (Approximate Differential Privacy [Dwork et al., 2006a])

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies (ϵ, δ) -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$

- For all measurable sets S , $P(M(\mathcal{D}_1) \in S) \leq e^\epsilon P(M(\mathcal{D}_2) \in S) + \delta$
- Or, equivalently, $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \geq \epsilon) - e^\epsilon P(\mathcal{L}_{\mathcal{D}_2, \mathcal{D}_1}^M \leq -\epsilon) \leq \delta$.


- : But what does it *mean*?

Example 2: Approximate Differential Privacy

Definition (Approximate Differential Privacy [Dwork et al., 2006a])

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies (ϵ, δ) -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$

- For all measurable sets S , $P(M(\mathcal{D}_1) \in S) \leq e^\epsilon P(M(\mathcal{D}_2) \in S) + \delta$
- Or, equivalently, $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \geq \epsilon) - e^\epsilon P(\mathcal{L}_{\mathcal{D}_2, \mathcal{D}_1}^M \leq -\epsilon) \leq \delta$.

• : But what does it *mean*?

- $\delta = 0 \Leftrightarrow$ pure differential privacy.
- $\epsilon = 0 \Leftrightarrow \sup_S |P(M(\mathcal{D}_1) \in S) - P(M(\mathcal{D}_2) \in S)| \leq \delta$
 - total variation distance
- ϵ is log odds
- δ ?

Example 2: Approximate Differential Privacy

Definition (Approximate Differential Privacy [Dwork et al., 2006a])

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies (ϵ, δ) -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$

- For all measurable sets S , $P(M(\mathcal{D}_1) \in S) \leq e^\epsilon P(M(\mathcal{D}_2) \in S) + \delta$
- Or, equivalently, $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \geq \epsilon) - e^\epsilon P(\mathcal{L}_{\mathcal{D}_2, \mathcal{D}_1}^M \leq -\epsilon) \leq \delta$.

• : But what does it *mean*?


- $\delta = 0 \Leftrightarrow$ pure differential privacy.
- $\epsilon = 0 \Leftrightarrow \sup_S |P(M(\mathcal{D}_1) \in S) - P(M(\mathcal{D}_2) \in S)| \leq \delta$
 - total variation distance
- ϵ is log odds
- δ ?
 - Probability ϵ -DP guarantees fail?
 - $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M > \epsilon) \leq \delta$?
 - **no**


Example 2: Approximate Differential Privacy

Definition (Approximate Differential Privacy [Dwork et al., 2006a])

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies (ϵ, δ) -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$

- For all measurable sets S , $P(M(\mathcal{D}_1) \in S) \leq e^\epsilon P(M(\mathcal{D}_2) \in S) + \delta$
- Or, equivalently, $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \geq \epsilon) - e^\epsilon P(\mathcal{L}_{\mathcal{D}_2, \mathcal{D}_1}^M \leq -\epsilon) \leq \delta$.

• : But what does it *mean*?

- $\delta = 0 \Leftrightarrow$ pure differential privacy.
- $\epsilon = 0 \Leftrightarrow \sup_S |P(M(\mathcal{D}_1) \in S) - P(M(\mathcal{D}_2) \in S)| \leq \delta$
 - total variation distance
- ϵ is log odds
- δ ?
 - $\delta \geq$ probability of catastrophic failure.
 - $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M = \infty) \leq \delta$.
 - This interpretation is useful applicable for most mechanisms.
 - Existence of intuitive interpretations unlikely.
 - Problem when asking policy makers to make decisions .
 - So what to do?

Attempt 1

- δ is not bound on $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M > \epsilon)$.
- So what if we redefine δ ?

Definition $((\epsilon, \delta)$ -Probabilistic Differential Privacy)

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies probabilistic DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$,

$$P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M > \epsilon) \leq \delta$$


Attempt 1

- δ is not bound on $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M > \epsilon)$.
- So what if we redefine δ ?

Definition ((ϵ, δ)-Probabilistic Differential Privacy)

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies probabilistic DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$,

$$P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M > \epsilon) \leq \delta$$

- (ϵ, δ)-curve
- δ more interpretable, specifies CDF of PLRV
- Probability of a “bad” ω (log odds higher than ϵ is $\leq \delta$).
- Probability that ϵ -DP guarantees fail is $\leq \delta$?
- Jupyter notebook example 


Attempt 1

- δ is not bound on $P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M > \epsilon)$.
- So what if we redefine δ ?

Definition ((ϵ, δ)-Probabilistic Differential Privacy)

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies probabilistic DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$,

$$P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M > \epsilon) \leq \delta$$

- (ϵ, δ)-curve
- δ more interpretable, specifies CDF of PLRV
- Probability of a “bad” ω (log odds higher than ϵ is $\leq \delta$).
- Probability that ϵ -DP guarantees fail is $\leq \delta$?
- Jupyter notebook example 
- **Bad news:** not postprocessing invariant.
- No adaptive composition properties

Attempt 2

- Lack of postprocessing invariance is easy to fix.
- Throw in conditions for $A \circ M$ for all postprocessing algorithms.

Attempt 2

- Lack of postprocessing invariance is easy to fix.
- Throw in conditions for $A \circ M$ for all postprocessing algorithms.

Definition ((ϵ, δ)-Probabilistically Bounded Differential Privacy (Census Working Paper))

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies pbdp if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$, and all (randomized) algorithms A whose domain contains the range of M ,

$$P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{A \circ M} > \epsilon) \leq \delta$$

Attempt 2

- Lack of postprocessing invariance is easy to fix.
- Throw in conditions for $A \circ M$ for all postprocessing algorithms.

Definition ((ϵ, δ)-Probabilistically Bounded Differential Privacy (Census Working Paper))

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies pbdp if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$, and all (randomized) algorithms A whose domain contains the range of M ,

$$P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^{A \circ M} > \epsilon) \leq \delta$$

- But this is so complex to work with

Attempt 2

- Lack of postprocessing invariance is easy to fix.
- Throw in conditions for $A \circ M$ for all postprocessing algorithms.

Definition ((ϵ, δ)-Probabilistically Bounded Differential Privacy (Census Working Paper))

Given $\epsilon > 0$ and $\delta \in (0, 1]$, M satisfies pbdp if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$, and all binary-valued (randomized) algorithms A whose domain contains the range of M ,

$$\left(P(A(M(\mathcal{D}_1)) = 1) \leq e^{-\epsilon} \delta \right) \Rightarrow \left(P(A(M(\mathcal{D}_2)) = 1) \leq \delta \right)$$

- Equivalent form
- ϵ : log odds
- δ : post-processing invariant upper bound on probability of large odds
 - A can be thought of an attacker
 - Or hypothesis test.

Attempt 2

- Lack of postprocessing invariance is easy to fix.
- Throw in conditions for $A \circ M$ for all postprocessing algorithms.

Definition ((ϵ, δ)-Probabilistically Bounded Differential Privacy (Census Working Paper))

Given $\epsilon > 0$ and $\delta \in (0, 1]$, M satisfies pbdp if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$, and all binary-valued (randomized) algorithms A whose domain contains the range of M ,

$$\left(P(A(M(\mathcal{D}_1)) = 1) \leq e^{-\epsilon} \delta \right) \Rightarrow \left(P(A(M(\mathcal{D}_2)) = 1) \leq \delta \right)$$

- Equivalent form
- ϵ : log odds
- δ : post-processing invariant upper bound on probability of large odds
 - A can be thought of an attacker
 - Or hypothesis test.
- **Bad news: not convex.**

The Curve

- Postprocessing fix: throw in conditions for $A \circ M$.

Definition ((ϵ, δ)-Probabilistically Bounded Differential Privacy (Census Working Paper))

Given $\epsilon > 0$ and $\delta \in (0, 1]$, M satisfies pbdp if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$, and all binary-valued (randomized) algorithms A whose domain contains the range of M ,

$$\left(P(A(M(\mathcal{D}_1))) = 1 \right) \leq e^{-\epsilon} \delta \Rightarrow \left(P(A(M(\mathcal{D}_2))) = 1 \right) \leq \delta$$

- Convexity fix: use a curve g or $f = 1 - g$

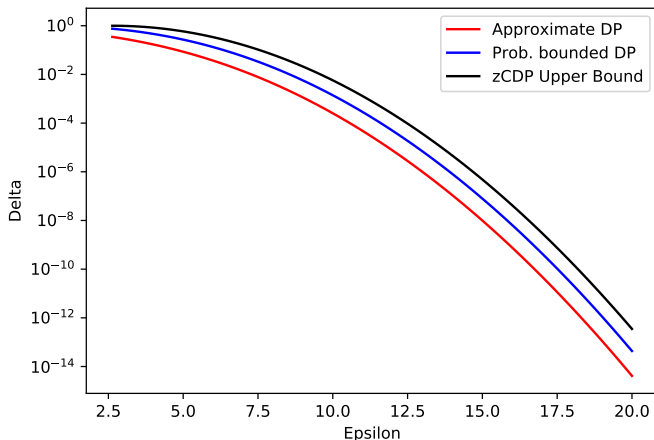
$$P(A(M(\mathcal{D}_2))) = 1 \leq g(P(A(M(\mathcal{D}_1))) = 1)$$

$$P(A(M(\mathcal{D}_2))) = 1 \leq 1 - f(P(A(M(\mathcal{D}_1))) = 1)$$

- $\epsilon = \frac{\delta}{g^{-1}(\delta)}$
- Conditions on f and g [Kifer and Lin, 2012, Dong et al., 2022]

(ϵ, δ) Curve

- Gaussian Mechanism with $N(0, \sigma^2 = \frac{1}{2 \cdot 2.63})$



f -DP

Definition (f -DP [Dong et al., 2022])

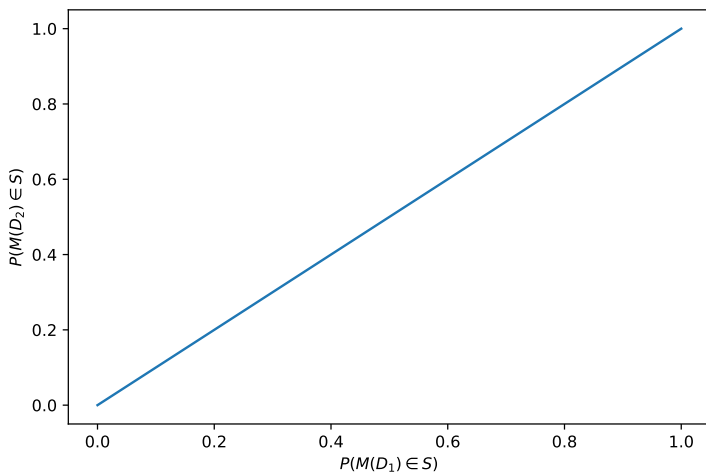
Let f be a continuous, convex, non-increasing function with $f(x) \leq 1 - x$. Mechanism M satisfies f -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$ and measurable S ,

$$P(M(\mathcal{D}_2) \in S) \leq 1 - f(P(M(\mathcal{D}_1) \in S))$$

- Census:
 - δ : (postprocessing invariant) probability that log odds are at least $\epsilon = \frac{\delta}{g^{-1}(\delta)}$, with $g = 1 - f$.
 - This is how δ is reported.
- Making δ in Approximate DP interpretable inevitably leads to f -DP.
- In practice, it makes sense to use f -DP.
 - Harder to work with.
 - But:
 - Better parameter interpretability
 - Good frequentist semantics
 - Good Bayesian semantics.

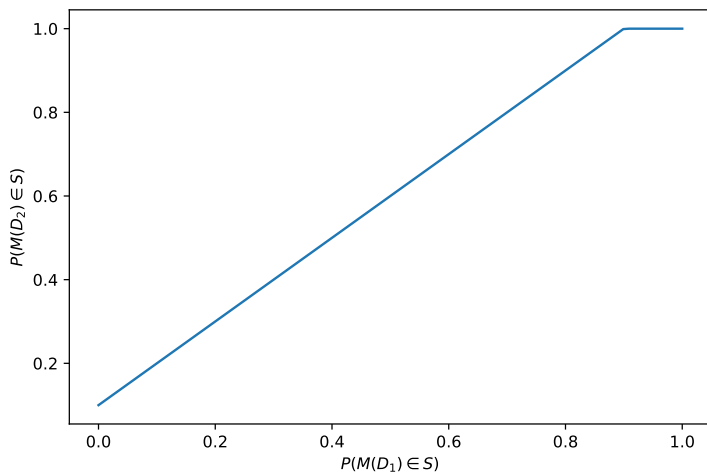
Tightest $1 - f$ curve.

- $M(\mathcal{D}) = 1$ (fully private)



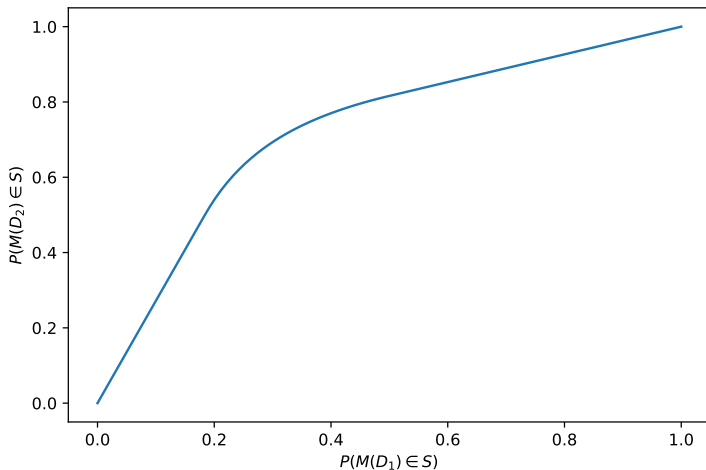
Tightest $1 - f$ curve.

- $M(\mathcal{D})$ = random record in dataset of 10 records



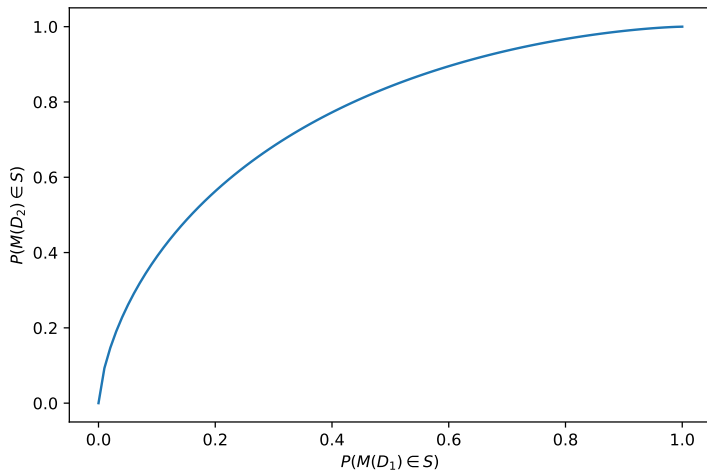
Tightest $1 - f$ curve.

- $M(\mathcal{D}) = \# \text{ AIAN} + \text{Laplace}(1)$



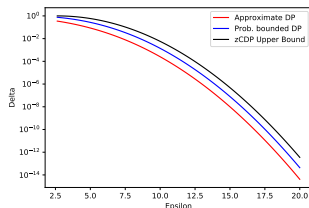
Tightest $1 - f$ curve.

- $M(\mathcal{D}) = \# \text{ AIAN} + N(0,1)$



Computation

- f function for a mechanism is often difficult to compute.
- Adaptive Composition is same as Independent Composition [Dong et al., 2022]
 - But independent composition is not always easy to calculate.
 - Currently no concept of what fully adaptive composition means for a function f .
- Rényi Differential Privacy (RDP) [Mironov, 2017] and zero-Concentrated Differential Privacy (zCDP) [Bun and Steinke, 2016] can help.
 - Privacy accounting frameworks.
 - Fully adaptive composition, simple composition rules.
 - Upper bounds on privacy parameters of other definitions.



RDP and zCDP

Definition ((α, γ)-Rényi Differential Privacy [Mironov, 2017])

Given an $\alpha > 1$, mechanism M satisfies (α, γ)-RDP if for all pairs of neighboring datasets $\mathcal{D}_1, \mathcal{D}_2$, $E[e^{(\alpha-1)\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M}] \leq e^{(\alpha-1)\gamma}$.

Definition (ρ -zero-Concentrated Differential Privacy [Bun and Steinke, 2016])

A mechanism M satisfies ρ -zCDP if for all pairs of neighboring datasets $\mathcal{D}_1, \mathcal{D}_2$ and all $\alpha > 1$, $E[e^{(\alpha-1)\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M}] \leq e^{\alpha(\alpha-1)\rho}$.

- Parameters not interpretable.
- Fully adaptive composition [Feldman and Zrnic, 2021]:
 - zCDP: ρ values add up.
 - RDP: if all α values are same, γ values add up.
- Gaussian Mechanism:
 - $\Delta_q = \sup_{\mathcal{D}_1, \mathcal{D}_2} \|q(\mathcal{D}_1) - q(\mathcal{D}_2)\|_2$
 - $q(D) + N(0, \sigma^2)$ is ρ -zCDP with $\rho = \frac{\Delta_q^2}{2\sigma^2}$, same with discrete Gaussian.

Privacy Definitions Summary

- Interpretability:
 - Pure differential privacy
 - f -DP/pbdp
- Frequentist semantics (next):
 - f -DP preferred
- Bayesian semantics (after):
 - Pure DP (most powerful)
 - f -DP (tightest)
 - RDP and zCDP (general)
- Computation and algorithm tuning:
 - Pure DP
 - RDP and zCDP (then convert)

Outline

- 1 The Big Picture
- 2 Privacy Loss Random Variables
- 3 From PLRV to Privacy Definitions
- 4 Frequentist Guarantees
- 5 Bayesian Guarantees
- 6 Application to the TopDown Algorithm
- 7 Utility and Uncertainty

Privacy Semantics

- “Kind of looks like DP” is not the same as “protects privacy”
- Privacy semantics: what is the effect on attacker’s inference.
- Explains to policy makers consequences of parameter choices.
- Philosophy + Math.

What is a violation of Privacy?



- Common beliefs:
 - Inference about me that is harmful.
 - Inference about me that would not be possible without this dataset.

What is a violation of Privacy?



- Common beliefs:
 - Inference about me that is harmful. ❌
 - Inference about me that would not be possible without this dataset. ❌
- It's deeper than most people realize.

When data are collected for statistical purposes.

- Produce data that supports generalizable (out-of-sample) statistical inference.
 - Voting surveys: voting pattern by demographics
 - Cancer studies: identify contributing factors to cancer
 - Demographic trends

When data are collected for statistical purposes.

- Produce data that supports generalizable (out-of-sample) statistical inference.
 - Voting surveys: voting pattern by demographics
 - Cancer studies: identify contributing factors to cancer
 - Demographic trends
- Statistical prediction \neq privacy breach
- Privacy breach: how you differ from the population
- Components of inference resulting from a data release:
 - Inference due to statistical information in data.
 - Inference due specifically to your information being in the data.
 - Provides information about how you differ from population.
 - Inclusion of record is a causal effect [Tschantz et al., 2020].

When data are collected for statistical purposes.

- Produce data that supports generalizable (out-of-sample) statistical inference.
 - Voting surveys: voting pattern by demographics
 - Cancer studies: identify contributing factors to cancer
 - Demographic trends
- Statistical prediction \neq privacy breach
- Privacy breach: how you differ from the population
- Components of inference resulting from a data release:
 - Inference due to statistical information in data.
 - Inference due specifically to your information being in the data.
 - Provides information about how you differ from population.
 - Inclusion of record is a causal effect [Tschantz et al., 2020].
- Belmont Report view of privacy:
 - Control over your information.
 - Perfect privacy: you withhold your data

DP View

- DP Semantics

- Actual world: your data is part of input to M .
- Counterfactual world: your data is scrubbed before it reaches M .
 - Privacy preserving baseline.
 - Eliminates causal effect of your record.

DP View

- DP Semantics
 - Actual world: your data is part of input to M .
 - Counterfactual world: your data is scrubbed before it reaches M .
 - Privacy preserving baseline.
 - Eliminates causal effect of your record.
- Alternative Prior-to-posterior view.
 - Belief about you prior to seeing the data should be almost same as after seeing the data.
 - Combines statistical knowledge and causal effect of your record.
 - Enforcing “posterior \approx prior” hurts data utility.

Frequentist View of DP Guarantees

- Frequentist semantics based on the output ω of M .
 - Null hypothesis H_0 : true data \mathcal{D}_1 is fed into M .
 - Alternative hypothesis H_1 : your record is first scrubbed, to create \mathcal{D}_2 , which is fed into M .
 - Counterfactual world.
 - Unbounded/add-remove neighbors: \mathcal{D}_2 is obtained by removing a record from \mathcal{D}_1 .
 - Bounded/modify-record neighbors: \mathcal{D}_2 is obtained by replacing a record with a “default” record.
 - Hypothesis test: any (randomized) algorithm A whose input is ω and output is 0 or 1.

Frequentist View of DP Guarantees

- Frequentist semantics based on the output ω of M .
 - Null hypothesis H_0 : true data \mathcal{D}_1 is fed into M .
 - Alternative hypothesis H_1 : your record is first scrubbed, to create \mathcal{D}_2 , which is fed into M .
 - Counterfactual world.
 - Unbounded/add-remove neighbors: \mathcal{D}_2 is obtained by removing a record from \mathcal{D}_1 .
 - Bounded/modify-record neighbors: \mathcal{D}_2 is obtained by replacing a record with a “default” record.
 - Hypothesis test: any (randomized) algorithm A whose input is ω and output is 0 or 1.
- Can attacker detect whether \mathcal{D}_1 or \mathcal{D}_2 was the input?
 - If yes: causal effect of including your record could be high.
 - If no: causal effect is small.
 - Semantics goal: attacker success is low for all neighbors $\mathcal{D}_1, \mathcal{D}_2$ (i.e., no data assumptions).
- Very compelling in practice for mathematical statisticians.

Hypothesis Test

- Null hypothesis $H_0 : \mathcal{D}_1$ (actual world)
- Alternative hypothesis $H_1 : \mathcal{D}_2$ (privacy preserving counterfactual world).
- ω is output of M .
- Hypothesis test A is a function of ω .
 - $A(\omega) = 1 \Rightarrow$ reject the null hypothesis (believes input is \mathcal{D}_2).
 - $A(\omega) = 0 \Rightarrow$ fail to reject null hypothesis (not enough evidence to distinguish between \mathcal{D}_1 and \mathcal{D}_2).
- Significance Level ℓ :
 - Type 1 error probability.
 - Reject null hypothesis when it is true.
 - $P(A(M(\mathcal{D}_1)) = 1)$.
- Power $1 - \beta$:
 - Reject null hypothesis when it is false.
 - $P(A(M(\mathcal{D}_2)) = 1)$
- If ω provides no information:
 - Hypothesis test is a random guess.
 - $\alpha = 1 - \beta$.

Differential Privacy

Definition

Given $\epsilon \geq 0$, M satisfies pure DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$ and all ω :

$$P(M(\mathcal{D}_1) = \omega) \leq e^\epsilon P(M(\mathcal{D}_2) = \omega)$$

- Postprocessing invariant.
- For any A , $P(A(M(\mathcal{D}_1)) = 1) \leq e^\epsilon P(A(M(\mathcal{D}_2)) = 1)$
- What restrictions on level ℓ and power $1 - \beta$ do we get?

Differential Privacy

Definition

Given $\epsilon \geq 0$, M satisfies pure DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$ and all ω :

$$P(M(\mathcal{D}_1) = \omega) \leq e^\epsilon P(M(\mathcal{D}_2) = \omega)$$

- Postprocessing invariant.
- For any A , $P(A(M(\mathcal{D}_1)) = 1) \leq e^\epsilon P(A(M(\mathcal{D}_2)) = 1)$
- What restrictions on level ℓ and power $1 - \beta$ do we get?
[Wasserman and Zhou, 2010]

$$\ell \leq e^\epsilon(1 - \beta)$$

$$1 - \beta \leq e^\epsilon \ell$$

$$(1 - \ell) \leq e^\epsilon \beta$$

$$\beta \leq e^\epsilon(1 - \ell)$$

$$\max(e^{-\epsilon} \ell, 1 - e^\epsilon(1 - \ell)) \leq 1 - \beta \leq \min(e^\epsilon \ell, 1 - e^{-\epsilon}(1 - \ell)).$$

Approximate Differential Privacy Guarantees?

Definition

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies (ϵ, δ) -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$ and measurable sets S ,

$$P(M(\mathcal{D}_1) \in S) \leq e^\epsilon P(M(\mathcal{D}_2) \in S) + \delta$$

Approximate Differential Privacy Guarantees?

Definition

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies (ϵ, δ) -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$ and measurable sets S ,

$$P(M(\mathcal{D}_1) \in S) \leq e^\epsilon P(M(\mathcal{D}_2) \in S) + \delta$$

[Kairouz et al., 2015]

$$\ell \leq e^\epsilon(1 - \beta) + \delta$$

$$1 - \beta \leq e^\epsilon \ell + \delta$$

$$(1 - \ell) \leq e^\epsilon \beta + \delta$$

$$\beta \leq e^\epsilon(1 - \ell) + \delta.$$

$$\begin{aligned} \max(e^{-\epsilon}(\ell - \delta), 1 - e^\epsilon(1 - \ell) - \delta) &\leq 1 - \beta \\ &\leq \min(e^\epsilon \ell + \delta, 1 - e^{-\epsilon}(1 - \ell - \delta)). \end{aligned}$$

Approximate Differential Privacy Guarantees?

Definition

Given $\epsilon \geq 0$ and $\delta \in [0, 1]$, M satisfies (ϵ, δ) -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$ and measurable sets S ,

$$P(M(\mathcal{D}_1) \in S) \leq e^\epsilon P(M(\mathcal{D}_2) \in S) + \delta$$

[Kairouz et al., 2015]

$$\ell \leq e^\epsilon(1 - \beta) + \delta$$

$$1 - \beta \leq e^\epsilon \ell + \delta$$

$$(1 - \ell) \leq e^\epsilon \beta + \delta$$

$$\beta \leq e^\epsilon(1 - \ell) + \delta.$$

$$\begin{aligned} \max(e^{-\epsilon}(\ell - \delta), 1 - e^\epsilon(1 - \ell) - \delta) &\leq 1 - \beta \\ &\leq \min(e^\epsilon \ell + \delta, 1 - e^{-\epsilon}(1 - \ell - \delta)). \end{aligned}$$

Each point on (ϵ, δ) curve adds more constraints.

Rényi Differential Privacy?

Definition

Given an $\alpha > 1$, mechanism M satisfies (α, γ) -RDP if for all pairs of neighboring datasets $\mathcal{D}_1, \mathcal{D}_2$,

$$\sum_{\omega} P(M(\mathcal{D}_1) = \omega)^\alpha P(M(\mathcal{D}_2) = \omega)^{\alpha-1} \leq e^{(\alpha-1)\gamma}$$

Semantics from [Balle et al., 2020].

zCDP

Definition

A mechanism M satisfies ρ -zCDP if for all pairs of neighboring datasets $\mathcal{D}_1, \mathcal{D}_2$ and all $\alpha > 1$,

$$\sum_{\omega} P(M(\mathcal{D}_1) = \omega)^{\alpha} P(M(\mathcal{D}_2) = \omega)^{\alpha-1} \leq e^{(\alpha-1)\alpha\rho}$$

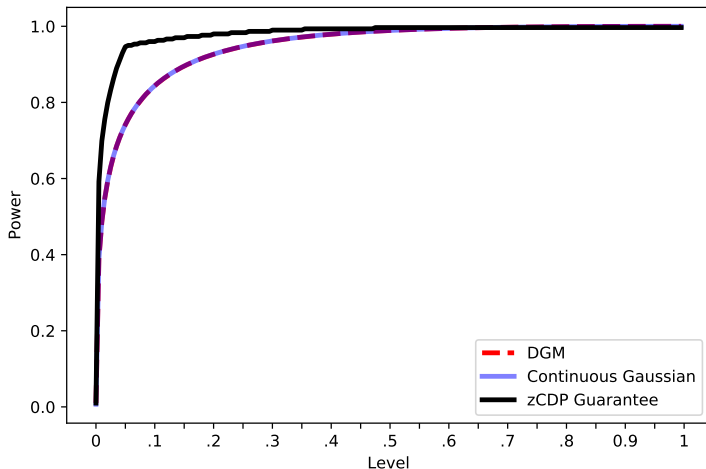
f -DP Frequentist Semantics?

Definition

Let f be a continuous, convex, non-increasing function with $f(x) \leq 1 - x$. Mechanism M satisfies f -DP if for all neighbors $\mathcal{D}_1, \mathcal{D}_2$ and measurable S ,

$$P(M(\mathcal{D}_2) \in S) \leq 1 - f(P(M(\mathcal{D}_1) \in S))$$

Redistricting Data Hypothesis Testing Semantics



Level (x-axis) vs. power (y-axis) curves for (1) the Gaussian mechanism at $\rho = 2.63$, (2) DGM, (3) arbitrary 2.63-zCDP algorithm.

Hypothesis Testing Semantics

Level	Power (Gaussian)	Power (DGM)	zCDP Upper Bound
0.01	0.49	0.49	0.70
0.05	0.74	0.74	0.95
0.10	0.84	0.84	0.96

Likelihood ratio test significance level/power trade-off for 2020 Census redistricting data production privacy-loss budget allocations (1) if Gaussian noise is used, (2) if discrete Gaussian noise is used, (3) guaranteed upper bound if an arbitrary ρ -zCDP mechanism with $\rho = 2.63$ is used.

Frequentist Semantics Recap

- Frequentist semantics can be read directly off the privacy definitions.
- If you can compute privacy parameters, you get the semantics.
- f -DP directly controls semantics.

Frequentist Semantics Recap

- Frequentist semantics can be read directly off the privacy definitions.
- If you can compute privacy parameters, you get the semantics.
- f -DP directly controls semantics.
- But what if you can't compute privacy parameters?
 - Discrete Gaussian Mechanism and f -DP
 - Upper bound (zCDP)
 - Sometimes you can estimate the significance level vs. power tradeoff.
 - Estimation error depends only on computation time.
 - Easier to do for non-adaptive algorithms.

Most Powerful Test

- Most powerful test of $H_0 : \mathcal{D}_1$ vs. $H_1 : \mathcal{D}_2$ is the likelihood ratio test.
 - If any test has the same significance level, likelihood ratio test has more power.
 - Defines upper boundary of the level vs. power curve.

Most Powerful Test

- Most powerful test of $H_0 : \mathcal{D}_1$ vs. $H_1 : \mathcal{D}_2$ is the likelihood ratio test.
 - If any test has the same significance level, likelihood ratio test has more power.
 - Defines upper boundary of the level vs. power curve.
- Test is defined by threshold t and tiebreaker c .
 - 1 Observe ω .
 - 2 For likelihood ratio: $r = \frac{P(M(\mathcal{D}_1)=\omega)}{P(M(\mathcal{D}_2)=\omega)}$
 - 3 If $r < t$, reject null hypothesis
 - 4 If $r = t$, reject null hypothesis with probability c .
 - 5 Otherwise fail to reject.

Most Powerful Test

- Most powerful test of $H_0 : \mathcal{D}_1$ vs. $H_1 : \mathcal{D}_2$ is the likelihood ratio test.
 - If any test has the same significance level, likelihood ratio test has more power.
 - Defines upper boundary of the level vs. power curve.
- Test is defined by threshold t and tiebreaker c .
 - 1 Observe ω .
 - 2 For likelihood ratio: $r = \frac{P(M(\mathcal{D}_1)=\omega)}{P(M(\mathcal{D}_2)=\omega)}$
 - 3 If $r < t$, reject null hypothesis
 - 4 If $r = t$, reject null hypothesis with probability c .
 - 5 Otherwise fail to reject.
- Significance level:
 - $(1 - c)P(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M < t) + cP(\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M \leq t)$
- Power:
 - $(1 - c)P(\mathcal{L}_{\mathcal{D}_2, \mathcal{D}_1}^M > -t) + cP(\mathcal{L}_{\mathcal{D}_2, \mathcal{D}_1}^M \geq -t)$
- Can be computed if we know CDFs of $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$ and $-\mathcal{L}_{\mathcal{D}_2, \mathcal{D}_1}^M$

Jupyter Notebook



Problem

- Guarantees have to hold for all neighboring pairs $\mathcal{D}_1, \mathcal{D}_2$.
- Need to compute level vs. power curve for all pairs.
- Take maximum curve (max power for each significance level).
- Generally intractable but algorithm structure can help.

Example: TDA simplification

- Geography levels: US, State, County, Tract, Block Group, Block.
- Queries:
 - e.g., q_1 : Male/Female histogram
 - e.g., q_2 : Voting Age/Non-Voting Age.
- Budget Allocation:
 - Global ρ .
 - Geography allocation: s_{usa}, s_{state} , that sum up to 1
 - Query allocation:
 - $v_{usa,1}$ for query 1 at national level
 - $v_{usa,2}$ for query 2 at national level
 - $v_{state,1}$ for query 1 at state level
- Query i at level j :
 - Budget allocation: $\tau = \rho * s_j * v_{j,i}$.
 - Noise added:
 - $N_Z(0, \sigma^2 = 1/\tau)$ for bounded neighbors
 - $N_Z(0, \sigma^2 = 1/(2 * \tau))$ for unbounded neighbors

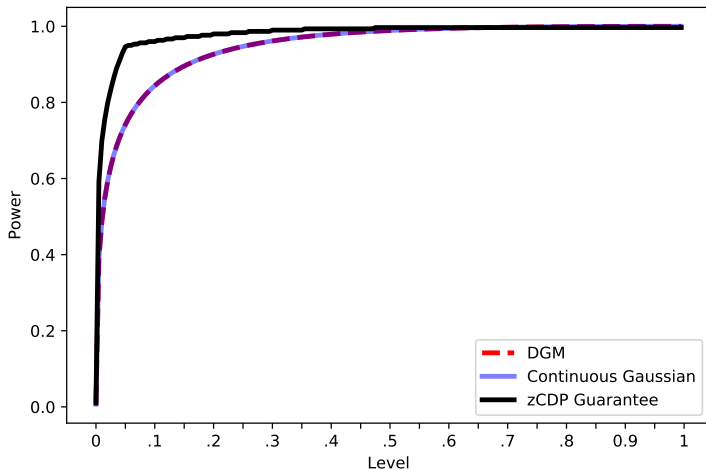
Example 1: no geographic hierarchy, 2 queries

- bounded neighbors.
- What are all the $\mathcal{L}_{\mathcal{D}_1, \mathcal{D}_2}^M$?
- What are the important ones (postprocessing can only reduce power)?
- Continuous vs. Discrete Gaussian.

Example 2: 2 queries, 2 levels

- Unbounded neighbors.
- Continuous vs. Discrete Gaussian.

Redistricting Data Hypothesis Testing Semantics



Level (x-axis) vs. power (y-axis) curves for (1) the Gaussian mechanism at $\rho = 2.63$, (2) DGM, (3) arbitrary 2.63-zCDP algorithm.

TDA Base ρ

Base ρ value for Person Characteristics and Housing Unit Characteristic, Production Settings [Abowd et al., html]

Person Characteristics	Housing Unit Characteristics
2.56	0.07

Geographic Allocation

Geographic Level	Person ρ Proportions	Housing Units ρ Prop.
US	104/4099	1/205
State	1440/4099	1/205
County	447/4099	7/82
Tract	687/4099	364/1025
Custom Block Group	1256/4099	1759/4100
Block	165/4099	99/820

Query Allocations

Query	Per Query ρ Allocation Proportions by Geographic Level					
	US	State	County	Tract	CBG	Block
TOTAL (1 cell)	0	3773/4097	3126/4097	1567/4102	1705/4099	5/4097
CENRACE (63 cells)	52/4097	6/4097	10/4097	4/2051	3/4099	9/4097
HISPANIC (2 cells)	26/4097	6/4097	10/4097	5/4102	3/4099	5/4097
VOTINGAGE (2 cells)	26/4097	6/4097	10/4097	5/4102	3/4099	5/4097
HHINSTLEVELS (3 cells)	26/4097	6/4097	10/4097	5/4102	3/4099	5/4097
HHGQ (8 cells)	26/4097	6/4097	10/4097	5/4102	3/4099	5/4097
HISPANIC \times CENRACE (126 cells)	130/4097	12/4097	28/4097	1933/4102	1055/4099	21/4097
VOTINGAGE \times CENRACE (126 cells)	130/4097	12/4097	28/4097	10/2051	9/4099	21/4097
VOTINGAGE \times HISPANIC (4 cells)	26/4097	6/4097	10/4097	5/4102	3/4099	5/4097
VOTINGAGE \times HISPANIC \times CENRACE (252 cells)	26/241	2/241	101/4097	67/4102	24/4099	71/4097
HHGQ \times VOTINGAGE \times HISPANIC \times CENRACE (2,016 cells)	189/241	230/4097	754/4097	241/2051	1288/4099	3945/4097

Outline

- 1 The Big Picture
- 2 Privacy Loss Random Variables
- 3 From PLRV to Privacy Definitions
- 4 Frequentist Guarantees
- 5 Bayesian Guarantees
- 6 Application to the TopDown Algorithm
- 7 Utility and Uncertainty

Bayesian Semantics

- Alternative statistical viewpoint.
- Used to model how attacker's background knowledge affects inference.
- Real world: actual data is run through a mechanism.
- Counterfactual world: target's data is first replaced by something else, then run through the mechanism.
 - Value of replaced record should not depend on actual record.
 - Value of replaced record does not depend on how actual record differs from statistical guesses.
- How similar is inference about target between actual and counterfactual world?

Bayesian Semantics

- Alternative statistical viewpoint.
- Used to model how attacker's background knowledge affects inference.
- Real world: actual data is run through a mechanism.
- Counterfactual world: target's data is first replaced by something else, then run through the mechanism.
 - Value of replaced record should not depend on actual record.
 - Value of replaced record does not depend on how actual record differs from statistical guesses.
- How similar is inference about target between actual and counterfactual world?
- Counterfactual world options (in bounded neighbor setting):
 - Target's record is replaced by a pre-chosen record.
 - "Perfectly private" (in the sense of Belmont Report) – target's data not used at all.

Warmup: Pure Differential Privacy

- Default record r_0
- Arbitrary prior $P_{\mathcal{Q}}$.
 - Probabilities that only depend on M are denoted using P .
 - Data random variable: \mathcal{X} .
 - Data without the target person: \mathcal{D}^- , \mathcal{X}^-

Warmup: Pure Differential Privacy

- Default record r_0
- Arbitrary prior $P_{\mathfrak{A}}$.
 - Probabilities that only depend on M are denoted using P .
 - Data random variable: \mathcal{X} .
 - Data without the target person: $\mathcal{D}^-, \mathcal{X}^-$
- Posterior-to-Posterior semantics: $\frac{P_{\mathfrak{A}}(r \mid M(\mathcal{X})=\omega)}{P_{\mathfrak{A}}(r \mid M(\mathcal{X}^- \cup \{r_0\})=\omega)}$
 [Kasiviswanathan and Smith, 2014]
 - Attacker is not trying to distinguish actual from counterfactual world
 - Attacker knows which world it is.
 - We are just comparing if the inference would be different.

Warmup: Pure Differential Privacy

- For **any** ω .
- Actual world posterior: $P_{\mathfrak{A}}(r \mid M(\mathcal{X}) = \omega)$

$$\frac{\sum_{\mathcal{D}^-} P_{\mathfrak{A}}(\mathcal{X}^- = \mathcal{D}^-) P_{\mathfrak{A}}(r \mid \mathcal{D}^-) P(M(\mathcal{D}^- \cup \{r\}) = \omega)}{\sum_{r'} \sum_{\mathcal{D}^-} P_{\mathfrak{A}}(\mathcal{X}^- = \mathcal{D}^-) P_{\mathfrak{A}}(r' \mid \mathcal{D}^-) P(M(\mathcal{D}^- \cup \{r'\}) = \omega)}$$

- Counterfactual world posterior:

$$\frac{\sum_{\mathcal{D}^-} P_{\mathfrak{A}}(\mathcal{X}^- = \mathcal{D}^-) P_{\mathfrak{A}}(r \mid \mathcal{D}^-) P(M(\mathcal{D}^- \cup \{r_0\}) = \omega)}{\sum_{r'} \sum_{\mathcal{D}^-} P_{\mathfrak{A}}(\mathcal{X}^- = \mathcal{D}^-) P_{\mathfrak{A}}(r' \mid \mathcal{D}^-) P(M(\mathcal{D}^- \cup \{r_0\}) = \omega)}$$

- Bound on ratio?

Bayesian Semantics

- Alternative statistical viewpoint.
- Used to model how attacker's background knowledge affects inference.
- Real world: actual data is run through a mechanism.
- Counterfactual world: target's data is first replaced by something else, then run through the mechanism.
 - Value of replaced record should not depend on actual record.
 - Value of replaced record does not depend on how actual record differs from statistical guesses.
- How similar is inference about target between actual and counterfactual world?
- Counterfactual world options (in bounded neighbor setting):
 - Target's record is replaced by a pre-chosen record.
 - Target's record is randomly sampled from a distribution.
 - Target's record is replaced by a random sample from $P_{\mathcal{A}}(r \mid \mathcal{D}^-)$.
 - "Perfectly private" (in the sense of Belmont Report) – target's data not used at all.

Semantics Pure Differential Privacy

- For **any** ω .
- Actual world posterior: $P_{\mathfrak{A}}(r \mid M(\mathcal{X}) = \omega)$

$$\frac{\sum_{\mathcal{D}^-} P_{\mathfrak{A}}(\mathcal{D}^-) P_{\mathfrak{A}}(r \mid \mathcal{D}^-) P(M(\mathcal{D}^- \cup \{r\}) = \omega)}{\sum_{r'} \sum_{\mathcal{D}^-} P_{\mathfrak{A}}(\mathcal{D}^-) P_{\mathfrak{A}}(r' \mid \mathcal{D}^-) P(M(\mathcal{D}^- \cup \{r'\}) = \omega)}$$

- Counterfactual world posterior:

$$\frac{\sum_{\mathcal{D}^-} P_{\mathfrak{A}}(\mathcal{D}^-) P_{\mathfrak{A}}(r \mid \mathcal{D}^-) \left(\sum_{r'} P_{\mathfrak{A}}(r' \mid \mathcal{D}^-) P(M(\mathcal{D}^- \cup \{r'\}) = \omega) \right)}{\sum_{r'} \sum_{\mathcal{D}^-} P_{\mathfrak{A}}(\mathcal{D}^-) P_{\mathfrak{A}}(r' \mid \mathcal{D}^-) P(M(\mathcal{D}^- \cup \{r'\}) = \omega)}$$

- Bound on ratio?

What about impure differential privacy?

- $\mathcal{D} = \underbrace{111 \cdots 111}_{100 \text{ bits}}$

$$M(\mathcal{D}) = \# \text{ of 1s} + N(0, 1).$$

What about impure differential privacy?

- $\mathcal{D} = \underbrace{111 \cdots 111}_{100 \text{ bits}}$

$$M(\mathcal{D}) = \# \text{ of 1s} + N(0, 1).$$

- Attacker prior:

- $P_{\mathfrak{A}}(\underbrace{100 \cdots 00}_{99 \text{ of these}}) = p = 0.01$

$$P_{\mathfrak{A}}(\underbrace{000 \cdots 00}_{99 \text{ of these}}) = 1 - p = 0.99$$

- Not an accurate prior.
 - Attacker believes target is independent of everyone else.

What about impure differential privacy?

- $\mathcal{D} = \underbrace{111 \cdots 111}_{100 \text{ bits}}$

$$M(\mathcal{D}) = \# \text{ of 1s} + N(0, 1).$$

- Attacker prior:

- $P_{\mathfrak{A}}(1 \underbrace{00 \cdots 00}_{99 \text{ of these}}) = p = 0.01$

$$P_{\mathfrak{A}}(0 \underbrace{00 \cdots 00}_{99 \text{ of these}}) = 1 - p = 0.99$$

- Not an accurate prior.
 - Attacker believes target is independent of everyone else.

- Typical output: $\omega = 101$

- Actual world posterior:
 - Counterfactual world posterior:

What did we learn?

- Pure DP protects against even perceived privacy breaches.
- Relaxations do not.

What did we learn?

- Pure DP protects against even perceived privacy breaches.
- Relaxations do not.
- Prior example shows:
 - Attacker has large difference between actual and counterfactual posteriors.
 - The ω attacker sees is practically impossible for the prior.
 - So rational attacker should abandon the prior.
- What kind of Bayesian guarantees are possible?

What did we learn?

- Pure DP protects against even perceived privacy breaches.
- Relaxations do not.
- Prior example shows:
 - Attacker has large difference between actual and counterfactual posteriors.
 - The ω attacker sees is practically impossible for the prior.
 - So rational attacker should abandon the prior.
- What kind of Bayesian guarantees are possible?
 - What if the attacker's prior is correct?
 - $\mathcal{D} \sim P_{\mathfrak{D}}(\mathcal{X})$
 - $\omega \sim M(\mathcal{D})$.
 - $\frac{P_{\mathfrak{D}}(r \mid M(\mathcal{X})=\omega)}{P_{\mathfrak{D}}(r \mid M(\text{counterfactual}(\mathcal{X}))=\omega)}$ is a random variable because ω is random.

What did we learn?

- Pure DP protects against even perceived privacy breaches.
- Relaxations do not.
- Prior example shows:
 - Attacker has large difference between actual and counterfactual posteriors.
 - The ω attacker sees is practically impossible for the prior.
 - So rational attacker should abandon the prior.
- What kind of Bayesian guarantees are possible?
 - What if the attacker's prior is correct?
 - $\mathcal{D} \sim P_{\mathfrak{A}}(\mathcal{X})$
 - $\omega \sim M(\mathcal{D})$.
 - $\frac{P_{\mathfrak{A}}(r \mid M(\mathcal{X})=\omega)}{P_{\mathfrak{A}}(r \mid M(\text{counterfactual}(\mathcal{X}))=\omega)}$ is a random variable because ω is random.
 - What is leaked about the true record r^* ?
 - $\mathcal{D} \sim P_{\mathfrak{A}}(\mathcal{X} \mid r^*)$
 - $\omega \sim M(\mathcal{D})$.
 - $\frac{P_{\mathfrak{A}}(r \mid M(\mathcal{X})=\omega)}{P_{\mathfrak{A}}(r \mid M(\text{counterfactual}(\mathcal{X}))=\omega)}$

The First Result [Kasiviswanathan and Smith, 2014]

- Dataset with n people
- M satisfies (ϵ, δ) -DP, $\delta < \epsilon^2/n$.
- ω is generated from M and $P_{\mathcal{X}}$
- Total variation distance:

$$\max_S |P_{\mathcal{X}}(\mathcal{X} \in S \mid \omega) - P_{\mathcal{X}}(\mathcal{X} \in S \mid \omega, \text{counterfactual})| < e^{3\epsilon} - 1 + 2\sqrt{n\delta}$$

- Except with probability at most $4\sqrt{n\delta}$

The First Result [Kasiviswanathan and Smith, 2014]

- Dataset with n people
- M satisfies (ϵ, δ) -DP, $\delta < \epsilon^2/n$.
- ω is generated from M and $P_{\mathcal{A}}$
- Total variation distance:

$$\max_S |P_{\mathcal{A}}(\mathcal{X} \in S \mid \omega) - P_{\mathcal{A}}(\mathcal{X} \in S \mid \omega, \text{counterfactual})| < e^{3\epsilon} - 1 + 2\sqrt{n\delta}$$

- Except with probability at most $4\sqrt{n\delta}$
- What we learn
 - If we don't specify a curve, δ should be much smaller than n .
 - $\epsilon < 0.23105$ for bound to be ≤ 1 .
 - Maybe (ϵ, δ) curve can help.

Census Bayesian Results

- Frequentist results appeal to many policy makers with statistical background.
- Bayesian results also provide value, more general questions can be asked:
 - Leakage about true record vs. leakage about a specific record.
 - Different assumptions about attacker knowledge.
 - Different comparison metrics.
 - Vast space to explore.

Semantics 1 (RDP, zCDP)

- Attacker knows \mathcal{D}^- , arbitrary $P_{\mathfrak{A}}(r \mid \mathcal{D}^-)$.
- For any fixed record:

Theorem

If M satisfies (α, γ) -RDP for $\alpha > 1$, then

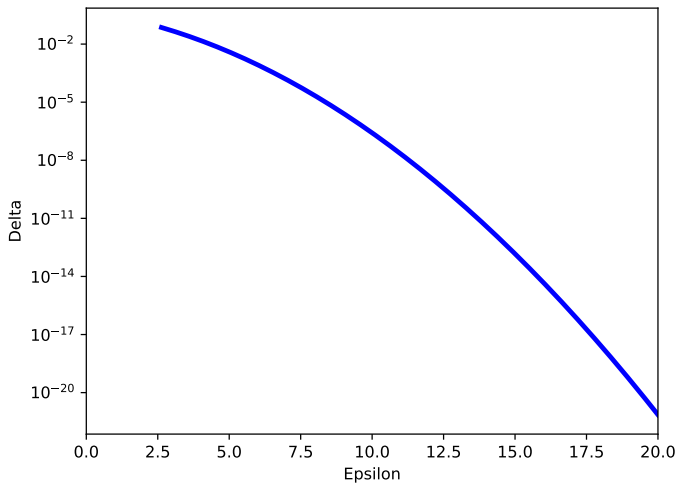
$$P_{\omega} \left(\frac{P_{\mathfrak{A}}(r \mid M(\mathcal{X})=\omega)}{P_{\mathfrak{A}}(r \mid M(\text{counterfactual}(\mathcal{X}))=\omega)} \right) \leq e^{-(\epsilon-\gamma)\alpha-\gamma}$$

Theorem

If M satisfies ρ -zCDP and $\epsilon > \rho$, then

$$P_{\omega} \left(\frac{P_{\mathfrak{A}}(r \mid M(\mathcal{X})=\omega)}{P_{\mathfrak{A}}(r \mid M(\text{counterfactual}(\mathcal{X}))=\omega)} \geq e^{\epsilon} \right) \leq e^{-(\epsilon+\rho)^2/(4\rho)}.$$

- Postprocessing invariant.
 - Can replace “ $M(\mathcal{X}) = \omega$ ” with “ $A(M(\mathcal{X})) \in S$ ”
 - Important property for semantics.

Semantics 1, $\rho = 2.63$ zCDP

Semantics 2 (PBDP)

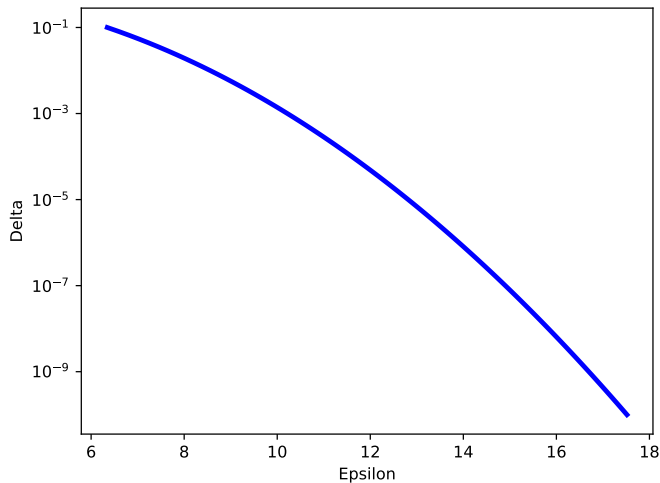
- Attacker knows \mathcal{D}^- , arbitrary $P_{\mathfrak{A}}(r \mid \mathcal{D}^-)$.
- Leakage about the true record:
 - $\mathcal{D} \sim P_{\mathfrak{A}}(\mathcal{X} \mid r^*)$
 - $\omega \sim M(\mathcal{D})$.

Theorem

If M satisfies an (ϵ, δ) -PBDP curve then

$$P_{\omega} \left(\frac{P_{\mathfrak{A}}(r \mid M(\mathcal{X})=\omega)}{P_{\mathfrak{A}}(r \mid M(\text{counterfactual}(\mathcal{X}))=\omega)} \right) \leq \delta$$

- Postprocessing invariant.
 - Can replace “ $M(\mathcal{X}) = \omega$ ” with “ $A(M(\mathcal{X})) \in S$ ”
 - Important property for semantics.
 - More support for f -DP.

Semantics 2, Gaussian Mechanism under $\rho = 2.63$ zCDP

Semantics 3 (RDP, zCDP)

- **Arbitrary** $P_{\mathfrak{A}}(r \mid \mathcal{D}^-)$
- Leakage about the true record:
 - $\mathcal{D} \sim P_{\mathfrak{A}}(\mathcal{X} \mid r^*)$
 - $\omega \sim M(\mathcal{D})$.

Theorem

If M satisfies (α, γ) -RDP for $\alpha > 1$, then

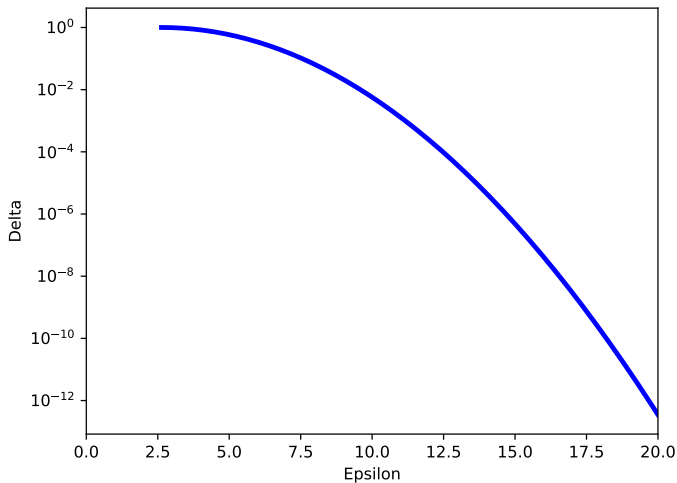
$$P_{\omega} \left(\frac{P_{\mathfrak{A}}(r \mid M(\mathcal{X})=\omega)}{P_{\mathfrak{A}}(r \mid M(\text{counterfactual}(\mathcal{X}))=\omega)} \right) \leq e^{(\alpha-1)(\gamma-\epsilon)}$$

Theorem

If M satisfies ρ -zCDP and $\epsilon > \rho$, then

$$P_{\omega} \left(\frac{P_{\mathfrak{A}}(r \mid M(\mathcal{X})=\omega)}{P_{\mathfrak{A}}(r \mid M(\text{counterfactual}(\mathcal{X}))=\omega)} \geq e^{\epsilon} \right) \leq e^{-(\epsilon-\rho)^2/(4\rho)}.$$

- Postprocessing invariant.
 - Can replace “ $M(\mathcal{X}) = \omega$ ” with “ $A(M(\mathcal{X})) \in S$ ”
 - Important property for semantics.

Semantics 3, $\rho = 2.63$ zCDP

Outline

- 1 The Big Picture
- 2 Privacy Loss Random Variables
- 3 From PLRV to Privacy Definitions
- 4 Frequentist Guarantees
- 5 Bayesian Guarantees
- 6 Application to the TopDown Algorithm
- 7 Utility and Uncertainty

Per-Attribute Semantics

- Specialized inference when person only cares about protecting certain attributes.
 - E.g., only race is sensitive.
 - Attacker has statistical information.
 - May know attributes other than race.
 - This may be predictive of race, but person may only care about protecting how they are different from this statistical baseline.
- $\mathcal{D}_1, \mathcal{D}_2$ differ only on those attributes.
- Only queries affected by those attributes are relevant.
- Only privacy budget allocated to those attributes is relevant.

TDA Base ρ

Base ρ value for Person Characteristics and Housing Unit Characteristic, Production Settings [Abowd et al., html]

Person Characteristics	Housing Unit Characteristics
2.56	0.07

Geographic Allocation

Geographic Level	Person ρ Proportions	Housing Units ρ Prop.
US	104/4099	1/205
State	1440/4099	1/205
County	447/4099	7/82
Tract	687/4099	364/1025
Custom Block Group	1256/4099	1759/4100
Block	165/4099	99/820

Query Allocations

Query	Per Query ρ Allocation Proportions by Geographic Level					
	US	State	County	Tract	CBG	Block
TOTAL (1 cell)	0	3773/4097	3126/4097	1567/4102	1705/4099	5/4097
CENRACE (63 cells)	52/4097	6/4097	10/4097	4/2051	3/4099	9/4097
HISPANIC (2 cells)	26/4097	6/4097	10/4097	5/4102	3/4099	5/4097
VOTINGAGE (2 cells)	26/4097	6/4097	10/4097	5/4102	3/4099	5/4097
HHINSTLEVELS (3 cells)	26/4097	6/4097	10/4097	5/4102	3/4099	5/4097
HHGQ (8 cells)	26/4097	6/4097	10/4097	5/4102	3/4099	5/4097
HISPANIC \times CENRACE (126 cells)	130/4097	12/4097	28/4097	1933/4102	1055/4099	21/4097
VOTINGAGE \times CENRACE (126 cells)	130/4097	12/4097	28/4097	10/2051	9/4099	21/4097
VOTINGAGE \times HISPANIC (4 cells)	26/4097	6/4097	10/4097	5/4102	3/4099	5/4097
VOTINGAGE \times HISPANIC \times CENRACE (252 cells)	26/241	2/241	101/4097	67/4102	24/4099	71/4097
HHGQ \times VOTINGAGE \times HISPANIC \times CENRACE (2,016 cells)	189/241	230/4097	754/4097	241/2051	1288/4099	3945/4097

Exercise Scenario 1: Block within Block Group

- Only cares about protecting fine-grained location.
 - Only interested in protecting the block within block group.
 - Not block group, tract, county, state.
 - Not race or any other demographic attribute.
 - Power vs. Level tradeoff?

Exercise Scenario 2: Block within tract?

Exercise Scenario 3: Block within blockgroup and race.

Outline

- 1 The Big Picture
- 2 Privacy Loss Random Variables
- 3 From PLRV to Privacy Definitions
- 4 Frequentist Guarantees
- 5 Bayesian Guarantees
- 6 Application to the TopDown Algorithm
- 7 Utility and Uncertainty

Privacy and Utility

- Data release:
 - Noisy query answers
 - Postprocessing
- Privacy is a multidimensional thing.
 - Many ways to think about it.
 - Do not optimize for a single (ϵ, δ) point.
- Utility: also complex.
- Topics:
 - Query selection: what to add noise to.
 - Challenges with privacy-protected microdata (why I am not a fan)

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 1: add noise to X and Y

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 1: add noise to X and Y
- Sensitivity:
 - For any database, adding/removing one person can
 - Change X by ± 1 .
 - Change Y by ± 1 .
 - Total change at most 2
 - Sensitivity Δ : 2

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 1: add noise to X and Y
- Sensitivity:
 - For any database, adding/removing one person can
 - Change X by ± 1 .
 - Change Y by ± 1 .
 - Total change at most 2
 - Sensitivity Δ : 2
- Noisy Counts (Measure):
 - $\tilde{X} = X + \text{Laplace}(2/\epsilon)$
 - $\tilde{Y} = Y + \text{Laplace}(2/\epsilon)$

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 1: add noise to X and Y
- Sensitivity:
 - For any database, adding/removing one person can
 - Change X by ± 1 .
 - Change Y by ± 1 .
 - Total change at most 2
 - Sensitivity Δ : 2
- Noisy Counts (Measure):
 - $\tilde{X} = X + \text{Laplace}(2/\epsilon)$
 - $\tilde{Y} = Y + \text{Laplace}(2/\epsilon)$
- Accuracy:
 - $\text{Var}(\tilde{X}) = 8/\epsilon^2$
 - $\text{Var}(\tilde{Y}) = 8/\epsilon^2$

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 2:
 - Add noise to $S = X + Y$ (Hispanic + VotingAge)
 - Add noise to $D = X - Y$ (Hispanic - VotingAge)
 - Note: not very intuitive quantities.

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 2:
 - Add noise to $S = X + Y$ (Hispanic + VotingAge)
 - Add noise to $D = X - Y$ (Hispanic - VotingAge)
 - Note: not very intuitive quantities.
- Sensitivity? In any database, adding/removing individual who is:
 - Neither Hispanic nor VotingAge: S and D unchanged.
 - Hispanic but not VotingAge: S changes by ± 1 , D changes by ± 1
 - Not Hispanic, is VotingAge: S changes by ± 1 , D changes by ± 1
 - Both Hispanic and VotingAge: S changes by ± 2 , D is unchanged.
 - Maximum change: 2
 - Sensitivity: 2

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 2:
 - Add noise to $S = X + Y$ (Hispanic + VotingAge)
 - Add noise to $D = X - Y$ (Hispanic - VotingAge)
 - Note: not very intuitive quantities.
- Sensitivity? In any database, adding/removing individual who is:
 - **Neither Hispanic nor VotingAge: S and D unchanged.**
 - Hispanic but not VotingAge: S changes by ± 1 , D changes by ± 1
 - Not Hispanic, is VotingAge: S changes by ± 1 , D changes by ± 1
 - Both Hispanic and VotingAge: S changes by ± 2 , D is unchanged.
 - Maximum change: 2
 - Sensitivity: 2

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 2:
 - Add noise to $S = X + Y$ (Hispanic + VotingAge)
 - Add noise to $D = X - Y$ (Hispanic - VotingAge)
 - Note: not very intuitive quantities.
- Sensitivity? In any database, adding/removing individual who is:
 - Neither Hispanic nor VotingAge: S and D unchanged.
 - **Hispanic but not VotingAge: S changes by ± 1 , D changes by ± 1**
 - Not Hispanic, is VotingAge: S changes by ± 1 , D changes by ± 1
 - Both Hispanic and VotingAge: S changes by ± 2 , D is unchanged.
 - Maximum change: 2
 - Sensitivity: 2

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 2:
 - Add noise to $S = X + Y$ (Hispanic + VotingAge)
 - Add noise to $D = X - Y$ (Hispanic - VotingAge)
 - Note: not very intuitive quantities.
- Sensitivity? In any database, adding/removing individual who is:
 - Neither Hispanic nor VotingAge: S and D unchanged.
 - Hispanic but not VotingAge: S changes by ± 1 , D changes by ± 1
 - **Not Hispanic, is VotingAge: S changes by ± 1 , D changes by ± 1**
 - Both Hispanic and VotingAge: S changes by ± 2 , D is unchanged.
 - Maximum change: 2
 - Sensitivity: 2

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 2:
 - Add noise to $S = X + Y$ (Hispanic + VotingAge)
 - Add noise to $D = X - Y$ (Hispanic - VotingAge)
 - Note: not very intuitive quantities.
- Sensitivity? In any database, adding/removing individual who is:
 - Neither Hispanic nor VotingAge: S and D unchanged.
 - Hispanic but not VotingAge: S changes by ± 1 , D changes by ± 1
 - Not Hispanic, is VotingAge: S changes by ± 1 , D changes by ± 1
 - **Both Hispanic and VotingAge: S changes by ± 2 , D is unchanged.**
 - Maximum change: 2
 - Sensitivity: 2

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 2:
 - Add noise to $S = X + Y$ (Hispanic + VotingAge)
 - Add noise to $D = X - Y$ (Hispanic - VotingAge)
 - Note: not very intuitive quantities.
- Sensitivity? In any database, adding/removing individual who is:
 - Neither Hispanic nor VotingAge: S and D unchanged.
 - Hispanic but not VotingAge: S changes by ± 1 , D changes by ± 1
 - Not Hispanic, is VotingAge: S changes by ± 1 , D changes by ± 1
 - Both Hispanic and VotingAge: S changes by ± 2 , D is unchanged.
 - Maximum change: 2
 - Sensitivity: 2

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 2:
 - Add noise to $S = X + Y$ (Hispanic + VotingAge)
 - Add noise to $D = X - Y$ (Hispanic - VotingAge)
 - Note: not very intuitive quantities.
- Sensitivity? Equals 2
- Noisy Measurements:
 - $\tilde{S} = S + \text{Laplace}(2/\epsilon)$
 - $\tilde{D} = D + \text{Laplace}(2/\epsilon)$

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 2:
 - Add noise to $S = X + Y$ (Hispanic + VotingAge)
 - Add noise to $D = X - Y$ (Hispanic - VotingAge)
 - Note: not very intuitive quantities.
- Sensitivity? Equals 2
- Noisy Measurements:
 - $\tilde{S} = S + \text{Laplace}(2/\epsilon)$
 - $\tilde{D} = D + \text{Laplace}(2/\epsilon)$
- Postprocess:
 - $\tilde{X} = (\tilde{S} + \tilde{D})/2$
 - $\tilde{Y} = (\tilde{S} - \tilde{D})/2$

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (Pure DP, unbounded neighbors)
- Attempt 2:
 - Add noise to $S = X + Y$ (Hispanic + VotingAge)
 - Add noise to $D = X - Y$ (Hispanic - VotingAge)
 - Note: not very intuitive quantities.
- Sensitivity? Equals 2
- Noisy Measurements:
 - $\tilde{S} = S + \text{Laplace}(2/\epsilon)$
 - $\tilde{D} = D + \text{Laplace}(2/\epsilon)$
- Postprocess:
 - $\tilde{X} = (\tilde{S} + \tilde{D})/2$
 - $\tilde{Y} = (\tilde{S} - \tilde{D})/2$
- Accuracy:
 - $\text{Var}(\tilde{S}) = 8/\epsilon^2$
 - $\text{Var}(\tilde{D}) = 8/\epsilon^2$
 - $\text{Var}(\tilde{X}) = 4/\epsilon^2$
 - $\text{Var}(\tilde{Y}) = 4/\epsilon^2$

Summary

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to?
- Attempt 1:
 - Add noise to X
 - Add noise to Y
 - Variance: $8/\epsilon^2$
- Attempt 2:
 - Add noise to $X + Y$
 - Add noise to $X - Y$
 - Postprocess (reconstruct)
 - Variance: $4/\epsilon^2$
- Select-Measure-Reconstruct Paradigm [Li et al., 2010a].
- What you want is not always what you should add noise to.

Strategy Example

- In a given region, suppose we are interested in:
 - $X = \#$ of Hispanic individuals
 - $Y = \#$ of VotingAge individuals
- What do we add noise to? (zCDP, unbounded neighbors)
 - L_2 sensitivity: $\Delta = \sup_{\mathcal{D}_1 \sim \mathcal{D}_2} \|q(\mathcal{D}_1) - q(\mathcal{D}_2)\|_2$
 - Gaussian Mechanism: $q(\mathcal{D}) + N(0, \sigma^2 = \frac{\Delta^2}{2\rho})$
- Which strategy is better?
 - 1 $X = \#$ of Hispanic individuals, $Y = \#$ of VotingAge individuals
 - 2 $X + Y$ and $X - Y$.

General Setup

- Dataset \mathcal{D} as a histogram \mathbf{x} .

$$\begin{pmatrix} \mathbf{x}[1] \\ \mathbf{x}[2] \\ \mathbf{x}[3] \\ \mathbf{x}[4] \end{pmatrix} = \begin{array}{lll} \text{number of people not adult, not Hispanic} \\ \text{number of people adult, not Hispanic} \\ \text{number of people not adult, Hispanic} \\ \text{number of people adult, Hispanic} \end{array}$$

- Query q is a vector, answer is $\mathbf{q} \cdot \mathbf{x}$
- Query matrix \mathbf{Q} , each row is a query. $\mathbf{Q}\mathbf{x}$ are the answers

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{array}{l} \text{number of people who are adult} \\ \text{number of people who are Hispanic} \\ \text{number of people} \end{array}$$

The Matrix Mechanism [Li et al., 2010b]

- Query matrix Q
 - Chosen using privacy loss budget (e.g., exponential mechanism)
 - Chosen in advance (provides ability to have per-attribute semantics).
- Strategy query matrix B
 - Measurement Mechanism: $M(\mathbf{x}) = B\mathbf{x} + \text{noise with covariance matrix } \Sigma$
 - Noisy answers: \mathbf{z}
 - Usually Σ is diagonal, but correlated noise also studied [Yuan et al., 2016, Xiao et al., 2021]
- Reconstruction (postprocessing)
 - Estimate $\tilde{\mathbf{x}}$ of \mathbf{x} .
 - OLS: $\tilde{\mathbf{x}} \leftarrow \arg \min_{\tilde{\mathbf{x}}} (B\tilde{\mathbf{x}} - \mathbf{z})^T \Sigma^{-1} (B\tilde{\mathbf{x}} - \mathbf{z})$
 - NNLS: $\tilde{\mathbf{x}} \leftarrow \arg \min_{\tilde{\mathbf{x}}} (B\tilde{\mathbf{x}} - \mathbf{z})^T \Sigma^{-1} (B\tilde{\mathbf{x}} - \mathbf{z}) \text{ s.t. } \tilde{\mathbf{x}} \geq 0$
 - Answer queries with $\tilde{\mathbf{x}}$ (i.e., $B\tilde{\mathbf{x}}$).

Jupyter Notebook

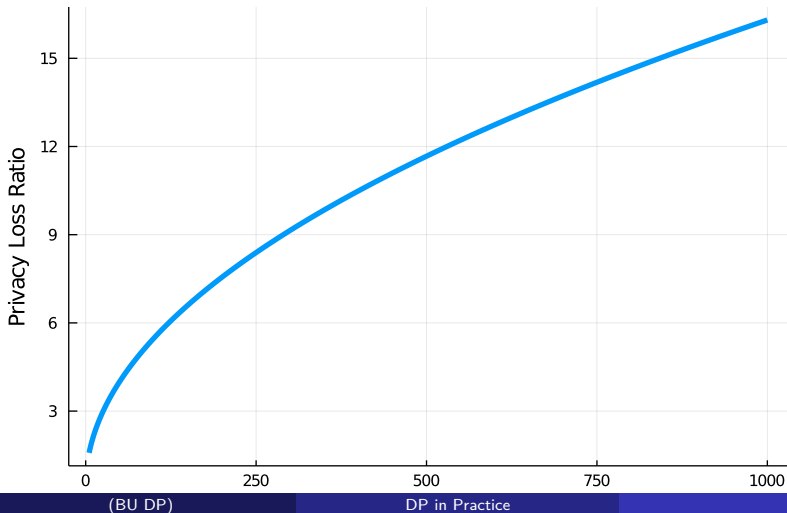


Fitness-for-Use

- In practice, privacy-protected output is used by many different users.
 - They don't care about total squared error.
 - They want to make sure queries they care about are fit-for-use.
 - "Accurate enough"
 - Many census data products use MOE.
 - Equivalent to variance constraints.
 - Problem: find strategy matrix B so that you hit every target constraint with minimal privacy cost [Xiao et al., 2021, Edmonds et al., 2020].
 - Not solved by the total squared error approach.
- Not the only possible notion of fitness for use.

Privacy loss sub-optimality

- Ratio of privacy loss of rescaled matrix mechanism to optimal mechanism.
- Privacy measured as in ρ -zCDP.



Jupyter Notebook



Challenges with privacy-protected microdata

- Many users want microdata
 - To use existing off-the-shelf tools
 - Often equivalent to ignoring noise in data.
- Dataset \mathcal{D} as a histogram \mathbf{x} .

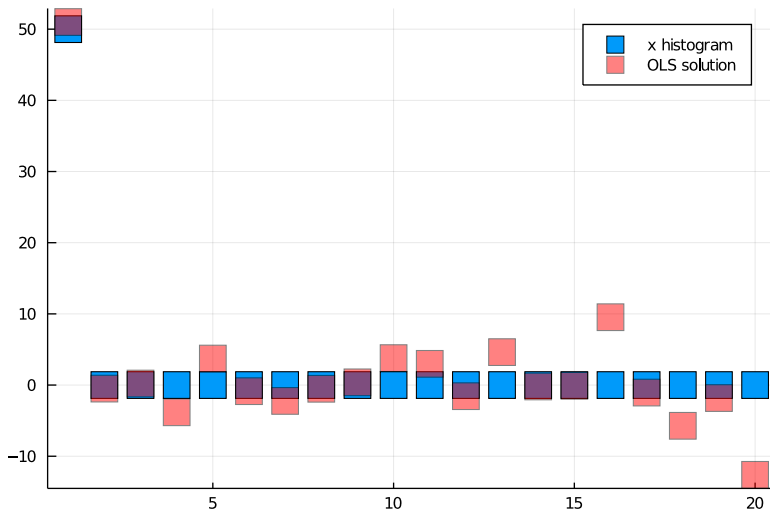
$$\begin{pmatrix} \mathbf{x}[1] \\ \mathbf{x}[2] \\ \mathbf{x}[3] \\ \mathbf{x}[4] \end{pmatrix} = \begin{array}{lll} \text{number of people not adult, not Hispanic} \\ \text{number of people} & \text{adult, not Hispanic} \\ \text{number of people not adult,} & \text{Hispanic} \\ \text{number of people} & \text{adult,} & \text{Hispanic} \end{array}$$

- Equivalent to microdata when \mathbf{x} is vector of nonnegative integers.
- Ignoring complications due to integers, let's see what happens when one query is the sum query and the others are the identity queries.
 - E.g., total county population for funding allocation (sum query).
 - E.g., population in each block in that county (redistricting).



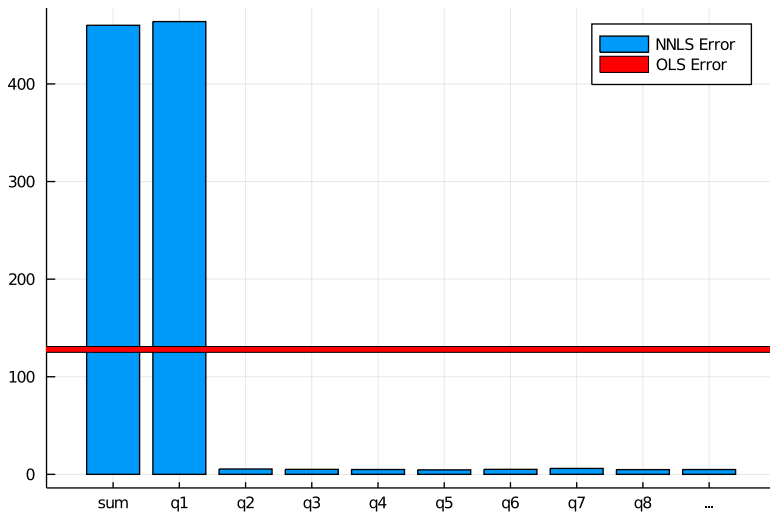
OLS Error

- Always at least as good as the noisy measurements (Gauss-Markov Theorem).



NNLS Error

- Often good, sometimes much worse than noisy measurements.



NNLS Observations

- Many privacy protected query answers become very accurate
- Some queries can get large systematic error. In toy example:
 - Sum query is inaccurate on average (not worst case)
 - First identity query is inaccurate on average (not worst case)
 - These are less accurate than noisy measurements we started with.
- Specifically because of nonnegativity.
- Not due to privacy budget allocation between queries.
- Is it avoidable?
 - Sometimes 🌫️
 - Exist datasets where this is unavoidable.

NNLS Observations

- Many privacy protected query answers become very accurate
- Some queries can get large systematic error. In toy example:
 - Sum query is inaccurate on average (not worst case)
 - First identity query is inaccurate on average (not worst case)
 - These are less accurate than noisy measurements we started with.
- Specifically because of nonnegativity.
- Not due to privacy budget allocation between queries.
- Your choice (no matter what DP algorithm you use):
 - Either guarantee sum query has error $O(1/\epsilon^2)$ and point queries may have **per-query** error up to $O(\log(n)^2/\epsilon^2)$.
 - Or guarantee point queries have error $O(1/\epsilon^2)$ but sum query has error up to $O(n^2/\epsilon^2)$.
 - Noisy measurements all have $O(1/\epsilon^2)$ error.
- Similar types of results for zCDP, approx DP
- Statistical price of microdata
[Balcer and Vadhan, 2019, Abowd et al., 2021]

Thank You



References I



Abowd, J., Ashmead, R., Cumings-Menon, R., Garfinkel, S., Kifer, D., Leclerc, P., Sexton, W., Simpson, A., Task, C., and Zhuravlev, P. (2021).

An uncertainty principle is a price of privacy-preserving microdata.
In NeurIPS.



Abowd, J. M., Ashmead, R., Cumings-Menon, R., Garfinkel, S., Heineck, M., Heiss, C., Johns, R., Kifer, D., Leclerc, P., Machanavajjhala, A., Moran, B., Sexton, W., Spence, M., and Zhuravlev, P. (forthcoming. Preprint <https://www.census.gov/library/working-papers/2022/adrm/CED-WP-2022-002.html>).

The 2020 census disclosure avoidance system topdown algorithm.
Harvard Data Science Review.

References II



Abowd, J. M. and Schmutte, I. M. (2015).
Economic analysis and statistical disclosure limitation.
Brookings Papers on Economic Activity, pages 221–267.



Alexander, J. T., Davern, M., and Stevenson, B. (2010).
Inaccurate age, and sex data in the census pums files: Evidence and
implications.
Public Opinion Quarterly, 74(3):551–69.



Balcer, V. and Vadhan, S. (2019).
Differential privacy on finite computers.
Journal of Privacy and Confidentiality, 9(2).

References III



Balle, B., Barthe, G., Gaboardi, M., Hsu, J., and Sato, T. (2020). Hypothesis testing interpretations and Rényi differential privacy. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*.



Bun, M. and Steinke, T. (2016). Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of Cryptography*.



Canonne, C. L., Kamath, G., and Steinke, T. (2020). The discrete gaussian for differential privacy. In *NeurIPS*.

References IV



Christ, M., Radway, S., and Bellovin, S. M. (2022).

Differential privacy and swapping: Examining de-identifications impact on minority representation and privacy preservation in the u.s. census.
In *SOSP*.



Dong, J., Roth, A., and Su, W. J. (2022).

Gaussian differential privacy.
JRSS-B, 84:3–37.



Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I., and Naor, M. (2006a).

Our data, ourselves: Privacy via distributed noise generation.
In Vaudenay, S., editor, *Advances in Cryptology - EUROCRYPT 2006*.



Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006b).

Calibrating noise to sensitivity in private data analysis.
In *Theory of cryptography conference*, pages 265–284. Springer.

References V



Edmonds, A., Nikolov, A., and Ullman, J. (2020).

The power of factorization mechanisms in local and central differential privacy.

In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 425–438.



Feldman, V. and Zrnic, T. (2021).

Individual privacy accounting via a Rényi filter.

In *Advances in Neural Information Processing Systems*.



Garfinkel, S. L. and Leclerc, P. (2020).

Randomness concerns when deploying differential privacy.

CoRR, abs/2009.03777.

References VI



Ilvento, C. (2020).

Implementing the Exponential Mechanism with Base-2 Differential Privacy, page 717742.

Association for Computing Machinery, New York, NY, USA.



Kairouz, P., Oh, S., and Viswanath, P. (2015).

The composition theorem for differential privacy.

In *Proceedings of the 32nd International Conference on Machine Learning (ICML)*.



Kasiviswanathan, S. P. and Smith, A. (2014).

On the 'Semantics' of Differential Privacy: A Bayesian Formulation.
Journal of Privacy and Confidentiality, 6(1):1–16.

See also: <https://arxiv.org/abs/0803.3946>.

References VII



Kifer, D. and Lin, B.-R. (2012).

An axiomatic view of statistical privacy and utility.

Journal of Privacy and Confidentiality, 4(1).



Li, C., Hay, M., Rastogi, V., Miklau, G., and McGregor, A. (2010a).

Optimizing linear counting queries under differential privacy.

In *PODS*.



Li, C., Hay, M., Rastogi, V., Miklau, G., and McGregor, A. (2010b).

Optimizing linear counting queries under differential privacy.

In *Proceedings of the Twenty-Ninth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*.

References VIII



Mironov, I. (2012).

On significance of the least significant bits for differential privacy.
In *Proceedings of the 2012 ACM Conference on Computer and Communications Security, CCS '12*, page 650661, New York, NY, USA. Association for Computing Machinery.



Mironov, I. (2017).

Rényi differential privacy.

In *30th IEEE Computer Security Foundations Symposium, CSF 2017, Santa Barbara, CA, USA, August 21-25, 2017*, pages 263–275.



Sommer, D. M., Meiser, S., and Mohammadi, E. (2019).

Privacy loss classes: The central limit theorem in differential privacy.
Proc. Priv. Enhancing Technol., 2019(2):245–269.

References IX



Tschantz, M., Sen, S., and Datta, A. (2020).

SoK: Differential privacy as a causal property.

In *2020 IEEE Symposium on Security and Privacy (SP)*, pages 354–371, Los Alamitos, CA, USA. IEEE Computer Society.



Warner, S. L. (1965).

Randomized response: A survey technique for eliminating evasive answer bias.

Journal of the American Statistical Association.



Wasserman, L. and Zhou, S. (2010).

A statistical framework for differential privacy.

Journal of the American Statistical Association, 105(489):375–389.



Xiao, Y., Ding, Z., Wang, Y., Zhang, D., and Kifer, D. (2021).

Optimizing fitness-for-use of differentially private linear queries.

Proc. VLDB Endow., 14(10).

References X



Xu, J., Zhang, Z., Xiao, X., Yang, Y., Yu, G., and Winslett, M. (2013).

Differentially private histogram publication.
The VLDB journal, 22(6):797–822.



Yuan, G., Yang, Y., Zhang, Z., and Hao, Z. (2016).

Convex optimization for linear query processing under approximate differential privacy.

In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.